Low-discrepancy Ensemble Kalman Filter

Uwe D. Hanebeck
Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
uwe.hanebeck@kit.edu

Ryne Beeson
Mechanical and Aerospace Engineering
Princeton University
Princeton, New Jersey, USA
ryne@princeton.edu

***Abstract*—We consider a specific class of filters, Ensemble Kalman Filters (EnKFs), that represent all occurring densities with particle ensembles. Unlike particle filters, EnKFs use Kalman filter-like updates. This has the benefit to avoid sample degeneration. Different variants of EnKFs exist to guarantee the posterior error covariance approximation produced by the updated ensemble matches that expected in the Kalman Filter (KF) setting. In this paper, we focus on improving the perturbed EnKF variant, where the predicted observations are perturbed with samples having the same statistics as the measurement noise. We propose a radically new type of EnKF with two contributions: (i) a specific type of low-discrepancy deterministic sampling is used for representing the noise and state densities that provide a much better coverage of observation and state spaces; and (ii) for enhanced prediction in both filter and prediction step, a new type of convolution of state ensembles with noise samples is devised, which improves performance especially for small ensemble sizes. Simulations show the increased estimation quality of the new Low-Discrepancy Ensemble Kalman Filter (LD-EnKF).**

I. Introduction

Estimating the latent state of a dynamic system from uncertain measurements is a ubiquitous problem in many applications ranging from robotics via autonomous driving to spacecraft navigation. Filtering is a special case of state estimation: Given a (noisy) continuous-time or discrete-time process model of the state evolution, a (noisy) model of the generation of measurements at a certain state for discrete time steps, and a set of (noisy) observations, the goal is to generate an optimal state estimate for each new observation. The uncertain state estimate is usually described by a probability density function (pdf) calculated recursively by alternating a prediction step based on the process model and a filter step based on the measurement model. The exact Bayesian Recursion Relations (BRR) are intractable for most interesting cases, so approximations are required.

A wealth of filtering methods exist that can be classified according to the type of approximate density representation employed and the various approximations of the prediction and filter step. Two common filter types are the Kalman Filter (KF) [20] and the Particle Filter (PF) [14]. The KF represents the state estimates by the first two moments and, crucially, assumes the state and measurement to be jointly Gaussian distributed. The PF is more general and represents the state densities by ensembles of weighted samples.

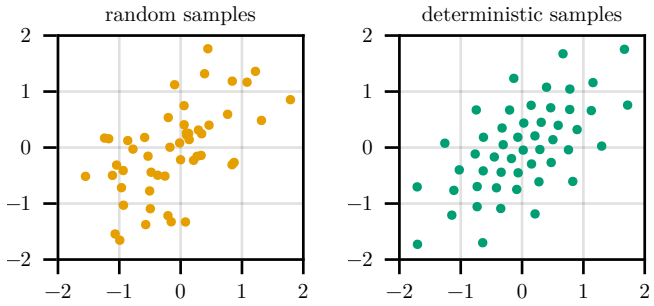

Fig. 1: Random versus deterministic ensemble for $L = 50$ samples.

The KF can be applied to nonlinear systems in various ways (see Sec. III), but due to its assumptions shows limited performance. In addition, calculation of the posterior covariance matrix is computationally complex and demands large memory storage requirements for high-dimensional problems. The PF typically produces higher quality-estimates, but is prone to particle degeneration, which is especially rapid in high-dimensional settings [3, 4, 31].

The EnKF is positioned between the KF and PF. It uses ensembles of particles for representing state densities like the PF and the same mechanism for the prediction step. However, in the filter step, instead of multiplying the ensemble members with the likelihood as in the PF, it uses simple Kalman filter-like updates applied to each ensemble member. This has the added benefit of avoiding the PF problem of sample degeneration. In the KF setting of linear models with additive Gaussian noise, the EnKF converges to the KF solution asymptotically for large ensemble size [7, 22, 25]. In the fully nonlinear and non-Gaussian setting, where the PF is able to provide a consistent solution to the Bayesian filtering problem in the asymptotic limit of number of particles, the EnKF does not provide a consistent result [22].

Due to its operating principle, the EnKF requires two predictions. The first occurs with the process model, where for discrete-time systems, the state at the current time step undergoes a generalized convolution with process noise samples to calculate the ensemble for the next time step. The second prediction occurs in the filter step, where measurements are predicted based on a generalized convolution of the current state ensemble with the measurement noise samples. These convolutions are performed by combining each member of

the state ensemble with a corresponding noise sample, thus maintaining the ensemble size.

There are two classes of EnKF methods that differ in the handling of the filter step. In this work, we focus on what is known as perturbed EnKFs, which are methods that use random samples, and will be detailed in Sec. V-A.

Random ensembles, especially for small ensemble sizes in high-dimensional spaces, are a poor representation of the underlying continuous densities, see Fig. 1 (left). Therefore, in this paper we use low-discrepancy samples. Specifically, we use the deterministic samples obtained by optimization from [17] that provide a homogeneous coverage of the underlying densities, see Fig. 1 (right).

In addition, we derive an advanced convolution technique. Instead of a one-to-one combination of state ensembles with noise ensembles, we combine each state sample with each noise sample, resulting in a larger temporary ensemble. In order to avoid increasing ensemble sizes, optimal down-sampling is applied to the temporary ensemble after each convolution to produce the final ensemble. This leads to a better representation quality especially for small ensembles.

The structure of the paper is as follows. The estimation problem is formulated in Sec. II. We then apply the KF approach to a nonlinear systems (without linearization) in Sec. III to get a general nonlinear Kalman filter from which later results are derived. The basic notation and ensemble setup of the standard EnKF begins Sec. IV. This is followed by the operating principle of the standard EnKF. The state-of-the-art in EnKFs, of which there are several varieties, is described in detail in Sec. V, with emphasis on the perturbed EnKF, which motivates the work of this paper. Sec. VI describes a deterministic low-discrepancy sampling mechanism and its use in our proposed low-discrepancy EnKF. Evaluations are shown in Sec. VII and conclusions are drawn in Sec. VIII. Lastly, remarks on the limitations of the proposed filter and an outlook to future work are summarized in Sec. IX.

Notation: Vectors are denoted by underlined lowercase letters, matrices by boldface uppercase letters. Random quantities are (additionally) denoted by boldface lowercase letters.

II. Problem Formulation

For the task of estimating a random state of a dynamic system, we are given: (i) a process model describing the time evolution of the state, (ii) a measurement model relating the observations to the state, and (iii) an initial state estimate.

Both process and measurement models come in different complexities. Here, we consider: (i) linear models with additive noise (`LinAdd`), (ii) nonlinear models with additive noise (`NonlinAdd`), and (iii) nonlinear models with a general noise structure (`NonlinNonadd`).

Process Model: The discrete-time process model at time steps $k \in \mathbb{N} \cup \{0\}$ describes the stochastic evolution equation for a random state vector $\underline{\boldsymbol{x}}_k \in \mathbb{R}^N$. In the simplest case, the state evolves according to a linear model characterized by matrices $\mathbf{A}_k$ with additive noise $\underline{\boldsymbol{\xi}}_k$

$$\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_k\, \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{\xi}}_k \quad (\texttt{LinAdd}). \tag{1}$$

A slightly more complex model is given by the nonlinear function $\underline{a}_k$ with additive noise

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k(\underline{\boldsymbol{x}}_k) + \underline{\boldsymbol{\xi}}_k \quad (\texttt{NonlinAdd}). \tag{2}$$

The most complex process model we consider is nonlinear with an arbitrary noise structure

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{\xi}}_k) \quad (\texttt{NonlinNonadd}). \tag{3}$$

The process noise $\underline{\boldsymbol{\xi}}_k$ could stem from an arbitrary distribution, but for the sake of simplicity of further developments in this paper, we assume zero-mean Gaussian noise with covariance matrix $\mathbf{Q}_k$ (i.e., $\underline{\boldsymbol{\xi}}_k \sim \mathrm{N}(\underline{0}, \mathbf{Q}_k)$).

Measurement Model: The measurement model relates the state $\underline{\boldsymbol{x}}_k$ to observations $\underline{\boldsymbol{y}}_k$. The simplest measurement model has a linear operator $\mathbf{H}_k$ with additive noise $\underline{\boldsymbol{v}}_k$

$$\underline{\boldsymbol{y}}_k = \mathbf{H}_k\, \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{v}}_k \quad (\texttt{LinAdd}). \tag{4}$$

A nonlinear measurement model with additive noise is

$$\underline{\boldsymbol{y}}_k = \underline{h}_k(\underline{\boldsymbol{x}}_k) + \underline{\boldsymbol{v}}_k \quad (\texttt{NonlinAdd}). \tag{5}$$

And a more complex model is a nonlinear measurement with arbitrary noise structure

$$\underline{\boldsymbol{y}}_k = \underline{h}_k(\underline{\boldsymbol{x}}_k\, \underline{\boldsymbol{v}}_k) \quad (\texttt{NonlinNonadd}). \tag{6}$$

In analogy to the process noise, the measurement noise is assumed to stem from a zero-mean Gaussian distribution with covariance matrix $\mathbf{R}_k$ (i.e., $\underline{\boldsymbol{v}}_k \sim \mathrm{N}(\underline{0}, \mathbf{R}_k)$).

Process and measurement noise $\underline{\boldsymbol{\xi}}_k$, $\underline{\boldsymbol{v}}_k$ are assumed to be white and independent, as well as independent of samples from the initial distribution.

Estimation Task: We assume that for $k = 0$, an initial state estimate is known from the initial distribution $f(\underline{\boldsymbol{x}}_0)$. It could for instance stem from a Gaussian density. However, we assume that only samples of this density are given. Additionally, we assume by convention that $\underline{y}_0 = \underline{0}$ is deterministic and provides no additional information at the initial time.

Our goal is then to solve the standard estimation task; that is, find an approximate representation of the filtering distribution $f(\underline{\boldsymbol{x}}_k \,|\, \underline{\boldsymbol{y}}_k, \underline{\boldsymbol{y}}_{k-1}, \ldots, \underline{\boldsymbol{y}}_1)$ and the predicted distribution $f(\underline{\boldsymbol{x}}_{k+1} \,|\, \underline{\boldsymbol{y}}_k, \underline{\boldsymbol{y}}_{k-1}, \ldots, \underline{\boldsymbol{y}}_1)$ for $k > 0$ in the form of sample sets (either random or deterministic).

Omission of time index k: To simplify the notation and especially, to omit the time index when possible, we use the random vector $\underline{\boldsymbol{x}}^e$ for describing the state estimate with the filtering distribution at time k. Similarly, the random vector $\underline{\boldsymbol{x}}^p$ describing the predicted state at time $k+1$ is associated with the predicted distribution.

III. Nonlinear Kalman Filter

In this section, we quickly review the most general version of a nonlinear KF for approximately filtering nonlinear systems. It will serve as the basis for deriving the standard EnKF and its proposed new low-discrepancy variant. The derivation will be for the most general system considered (`NonlinNonadd`), and in particular focus on the measurement update. We will

assume in this section that a prior estimate $\underline{\boldsymbol{x}}^p$ with density $f^p(\underline{\boldsymbol{x}})$ is given. The prior could have been generated by any of the model equations (1), (2), or (3).

A. Filter Step (Measurement Update)

For deriving the filter step of the nonlinear Kalman filter, we additionally assume to be given: (i) a noise $\underline{\boldsymbol{v}}$ with density $f^v(\underline{\boldsymbol{v}})$, and (ii) the measurement model in (6). The KF assumes the prior and the measurement to be *jointly Gaussian distributed*, so we consider the joint vector $\underline{\boldsymbol{z}}$ given by stacking $\underline{\boldsymbol{x}}^p$ and the yet unknown measurement $\underline{\boldsymbol{y}}$ as

$$\underline{\boldsymbol{z}} = \begin{bmatrix} \underline{\boldsymbol{x}}^p \\ \underline{\boldsymbol{y}} \end{bmatrix}. \tag{7}$$

The expected value of $\underline{\boldsymbol{z}}$ is given by

$$\hat{\underline{z}} = \mathrm{E}\{\underline{\boldsymbol{z}}\} = \begin{bmatrix} \hat{\underline{x}}^p \\ \hat{\underline{y}} \end{bmatrix} \tag{8}$$

with

$$\hat{\underline{x}}^p = \mathrm{E}\{\underline{\boldsymbol{x}}^p\} \text{ and } \hat{\underline{y}} = \mathrm{E}\{\underline{h}(\underline{\boldsymbol{x}}^p, \underline{\boldsymbol{v}})\}. \tag{9}$$

The covariance matrix of $\underline{\boldsymbol{z}}$ is

$$\mathbf{C}^z = \begin{bmatrix} \mathbf{C}^p & \mathbf{C}^{py} \\ \mathbf{C}^{yp} & \mathbf{C}^y \end{bmatrix} \tag{10}$$

with

$$\mathbf{C}^p = \mathrm{E}\left\{(\underline{\boldsymbol{x}}^p - \hat{\underline{x}}^p)(\underline{\boldsymbol{x}}^p - \hat{\underline{x}}^p)^T\right\}, \tag{11}$$

$$\mathbf{C}^{py} = \mathrm{E}\left\{(\underline{\boldsymbol{x}}^p - \hat{\underline{x}}^p)(\underline{h}(\underline{\boldsymbol{x}}^p, \underline{\boldsymbol{v}}) - \hat{\underline{y}})^T\right\}, \tag{12}$$

$$\mathbf{C}^y = \mathrm{E}\left\{(\underline{h}(\underline{\boldsymbol{x}}^p, \underline{\boldsymbol{v}}) - \hat{\underline{y}})(\underline{h}(\underline{\boldsymbol{x}}^p, \underline{\boldsymbol{v}}) - \hat{\underline{y}})^T\right\}, \tag{13}$$

and $\mathbf{C}^{yp} = (\mathbf{C}^{py})^T$. $\mathbf{C}^p$ is the prior error covariance of the state. This Gaussian density can easily be conditioned on the observed measurement $\underline{y}$, which gives the posterior expected value

$$\hat{\underline{x}}^e = \hat{\underline{x}}^p + \mathbf{K}(\underline{y} - \hat{\underline{y}}) \tag{14}$$

with Kalman gain

$$\mathbf{K} = \mathbf{C}^{py}\,(\mathbf{C}^y)^{-1} \tag{15}$$

and posterior covariance matrix

$$\mathbf{C}^e = \mathbf{C}^p - \mathbf{C}^{py}\,(\mathbf{C}^y)^{-1}\,\mathbf{C}^{yp}. \tag{16}$$

The calculation of the matrix $\mathbf{C}^e$ is intractable in high-dimensional spaces due to the exponential increase in computation and storage from: (i) matrix inversion of $\mathbf{C}^y$, and (ii) the storage, formation and product of covariance matrices. Yet if $\mathbf{K}$ can be approximated and treated as a constant parameter in (14), then (14) is itself manageable in higher dimensions as a function of $\hat{\underline{x}}^p$ and $\hat{\underline{y}}$. This is the fact exploited by the EnKF (c.f., Sec. IV and (26) in particular).

B. Prediction Step (Time Update)

The prediction of the nonlinear Kalman filter can be derived analogously to the filter step. Given an estimate $\underline{\boldsymbol{x}}^e$ with mean $\hat{\underline{x}}^e$ and covariance matrix $\mathbf{C}^e$, process noise $\underline{\boldsymbol{\xi}}$ with given density, and the system model in (3), the predicted state $\underline{\boldsymbol{x}}^p$ at the next time step can be generated by calculating the appropriate expected values (and again serve as an input to the filter step in the previous subsection).

IV. Ensemble Kalman Filter

We now give a concise recapitulation of the operating principles for the standard EnKF, which is also referred to as a perturbed EnKF when it is necessary to delineate it from a second class of EnKF, the deterministic EnKF, that will be briefly highlighted in Sec. V. Unlike the standard approach for derivation of the EnKF, which assumes the (`LinAdd`) case of (4), here we maintain maximal generality by assuming the (`NonlinNonadd`) setup of (6).

A. Weighted Ensemble Representation

The distribution for a random variable $\underline{\boldsymbol{x}}$ of a state can be represented by a Dirac mixture distribution of the form

$$f(x) = \sum_{i=1}^{L} w_i\,\delta(\underline{x} - \underline{x}_i) \tag{17}$$

with $L \in \mathbb{N}$ components, positive weights w_i summing up to one, and sample locations $\underline{x}_i$. The weights are collected in a weight vector

$$\underline{w} = \left[w_1, w_2, \cdots, w_L\right] \tag{18}$$

and the samples are collected in a matrix

$$\mathbf{X} = \left[\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_L\right]. \tag{19}$$

It is generally preferable to work with equally weighted samples (i.e., $w_i = 1/L$), which is in particular true for independent samples. In vector-matrix notation, the empirical mean vector is then given by

$$\hat{\underline{x}} = \mathbf{X}\,\underline{w}. \tag{20}$$

With $\underline{1}_L$ defining a vector with length L and containing a one in each entry of the vector, the empirical covariance matrix is defined as

$$\mathbf{C} = c\ \mathrm{diag}(\underline{w})\,(\mathbf{X} - \hat{\underline{x}}\,\underline{1}_L^T)\,(\mathbf{X} - \hat{\underline{x}}\,\underline{1}_L^T)^T, \tag{21}$$

where $c = 1$ for deterministic samples and otherwise typically selected as $c = L/(L-1)$ for random samples to ensure an unbiased estimate of the covariance matrix.

B. Filter Step (Measurement Update)

The EnKF assumes that an ensemble of samples of the prior density $f^p(\underline{x})$ and samples of the noise density $f^v(\underline{v})$ are given. With these samples, the posterior mean vector in (14) and the posterior covariance matrix in (16) can be calculated. As mentioned at the end of Sec. III-A, the calculation of $\mathbf{C}^e$ is expensive, and therefore the key idea of the EnKF is to apply (14) to each prior ensemble member individually with a fixed $\mathbf{K}$ that needs to be approximated. The general filter step of the EnKF proceeds as follows.

The prior ensemble is defined by L samples with weights

$$\underline{w}^p = \left[w_1^p, w_2^p, \cdots, w_L^p\right] \tag{22}$$

and locations

$$\mathbf{X}^p = \left[\underline{x}_1^p, \underline{x}_2^p, \cdots, \underline{x}_L^p\right]. \tag{23}$$

In addition, consider an observational noise ensemble with L samples, with weights

$$\underline{w}^v = \left[w_1^v, w_2^v, \cdots, w_L^v\right], \tag{24}$$

and locations

$$\mathbf{V} = \left[\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_L\right]. \tag{25}$$

Each noise sample $\underline{v}_i$ is associated with the prior sample $\underline{x}_i^p$ with the same index i.

Then observations for each prior ensemble and its associated observation noise can be calculated as

$$\hat{\underline{y}}_i = \underline{h}(\underline{x}_i^p, \underline{v}_i), \quad \text{for} \quad i \in [1, 2, \ldots, L]. \tag{26}$$

The Kalman gain (15) can be calculated using empirical covariances based on the given ensembles $\mathbf{X}^p$ and $\mathbf{V}$ so that the corresponding posterior ensemble $\mathbf{X}^e$ (with weights $\underline{w}^e = \underline{w}^p$) is defined by the mean recursion of (14) as

$$\underline{x}_i^e = \underline{x}_i^p + \mathbf{K}(\underline{y} - \hat{\underline{y}}_i), \qquad i \in [1, 2, \ldots, L]. \tag{27}$$

It is important to note that for the linear case (`LinAdd`), the empirical covariance of the posterior ensemble with samples given by (27) converges asymptotically in L to the KF covariance [12] and the empirical mean converges to the KF solution [7, 22, 25].

C. Prediction Step (Time Update)

For a given estimate represented by the ensemble $\mathbf{X}^e$ of size L at the current time step and a noise ensemble $\boldsymbol{\Xi}$ of size L, the predicted ensemble $\mathbf{X}^p$ for the next time step is obtained by combining each member of $\mathbf{X}^e$ with the corresponding member of $\boldsymbol{\Xi}$ based on the process model (3). The resulting ensemble $\mathbf{X}^p$ is also of size L and can be used as an input to the next filter step according to the previous subsection.

V. State-of-the-Art

The EnKF was first introduced in the geophysical sciences [11], shortly after the introduction of the PF [14]. Due to the enormous computational complexity incurred in solving high-dimensional geophysical problems, the EnKF is intentionally limited in the number of ensemble members used and therefore the measurement update of the EnKF settles for approximating the first two central moments of the posterior distribution using a Kalman-type update. In this way, the EnKF provides a closure condition for the nonlinear KF and also provides distribution representation improvement over variational methods [8, 32], which solve an expected maximum likelihood problem and are commonly used in the geosciences.

The typical derivation setting of the EnKF is the case (`LinAdd`), and the original equivalent of (27) in this setting was derived [11] as

$$\underline{x}_i^e = \underline{x}_i^p + \mathbf{K}(\underline{y} - \mathbf{H}\underline{x}_i^p), \qquad i \in [1, 2, \ldots, L]. \tag{28}$$

For the KF, the posterior error covariance $\mathbf{C}^e$ is updated (in the Joseph form) according to

$$\mathbf{C}^e = (\mathbf{I} - \mathbf{KH})\,\mathbf{C}^p\,(\mathbf{I} - \mathbf{KH})^T + \mathbf{KRK}^T, \tag{29}$$

where $\mathbf{I}$ is an identity matrix. It was mathematically articulated [6] and numerically observed [18] that although (28) provides the correct update for the empirical mean to match that of the KF, the ensemble posterior covariance does not via (28) match (29). It turns out that the ensemble covariance $\mathbf{C}^e$ will be reduced too much (i.e., is overconfident), which is a result of the ensemble covariance update missing the equivalent of the $\mathbf{KRK}^T$ term in (29). Corrections to address this issue resulted in the creation of two classes of EnKFs: the perturbed EnKF described in Sec. V-A and the deterministic EnKF briefly covered in Sec. V-B. The work in this manuscript is a new and novel approach to the class of perturbed EnKF. Some final remarks on standard modifications and philosophy of the EnKFs varieties, as well as review articles is given in Sec. V-C.

A. Perturbed EnKF

An ensemble approximation of the missing term $\mathbf{KRK}^T$ for the ensemble posterior covariance can be recovered by adding a random noise perturbation to the ensemble observation in the innovation term of (28) as

$$\underline{x}_i^e = \underline{x}_i^p + \mathbf{K}(\underline{y} - \mathbf{H}\underline{x}_i^p - \underline{v}_i), \qquad i \in [1, 2, \ldots, L]. \tag{30}$$

This was suggested by [6, 18], and in the asymptotic limit of L, the ensemble posterior covariance converges to the KF covariance [12]. Any EnKF approach that introduces random samples to the innovation in this manner is known as a perturbed EnKF (or sometimes as a stochastic EnKF).

In the (`LinAdd`) case with zero-mean noise, the end result is equivalent to viewing the random perturbations $\underline{v}_i$ as those of the given observation $\underline{y}$, and hence much of the literature talks about perturbed observations instead of the mathematically correct view of perturbed ensemble observations [34].

A key issue with the EnKF is that small ensemble sizes lead to spurious correlation errors. Therefore, the introduction of a low sample size of ensemble observation noise samples, that are paired one-to-one with the state ensembles, causes further spurious correlation errors. To attempt to reduce the correlation errors, improved sampling schemes were introduced [27, 10]. In the case of perturbed ensemble observation noises, these methods work by oversampling the observation noise (i.e., more than L samples), and then performing a matrix decomposition (e.g., a singular value decomposition) to generate a random truncated basis with the same rank as the ensemble size L that forms a matrix with an improved condition number.

Additional corrections can be made to the random samples, such as translating the mean and rescaling to get the correct covariance. Although these additional steps are unprincipled, they do show empirical improvement in reducing error in estimating the hidden state in general.

B. Deterministic EnKF

The second class of EnKF methods are known as deterministic EnKFs and are based on finding a square root update for the ensemble $\mathbf{X}^p$ such that the condition of (29) is exactly satisfied. This was first suggested in a set of papers by [2, 5, 36], all of which are reviewed with a uniform notation in

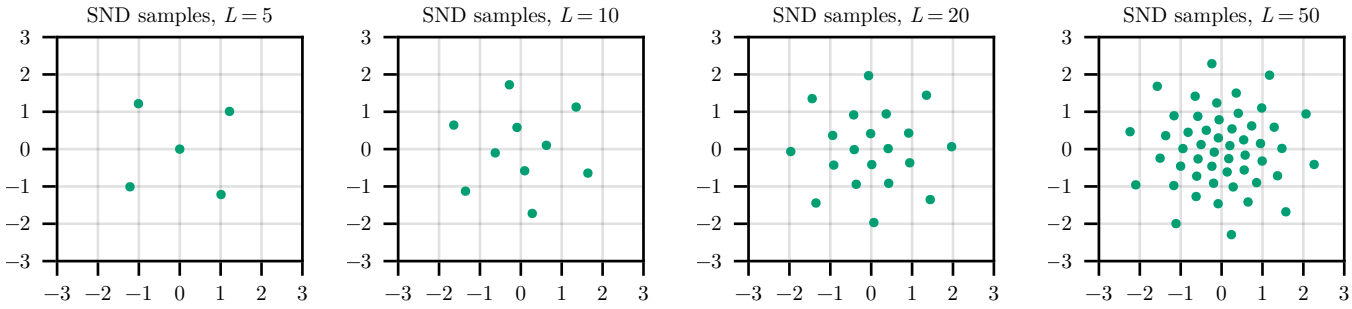

Fig. 2: Sets of deterministic samples representing a two-dimensional standard normal distribution (SND) for $L = 5, 10, 20$, and 50 samples.

[33]. More recent reviews appear in [26, 35], as well as a review aimed toward the signal processing community is [28]. Some studies have shown that for linear to weakly nonlinear problems with small ensemble sizes that deterministic EnKFs are superior to those described in Sec. V-A [12, 21] But for the truly nonlinear cases or moderate to higher ensemble sizes, a perturbed EnKF is likely preferred [21] and hence the focus of improvement for this manuscript.

C. Additional Remarks

Covariance Inflation and Localization: The low sample size of all EnKF methods relative to typical high-dimensional applications means that spurious correlations are still an issue, regardless of whether improved sampling schemes with a perturbed EnKF or a deterministic EnKF is used. It is therefore standard to apply covariance inflation and localization techniques in real high-dimensional problems. We refer the reader to [19] for a good overview of this subject. The work in this manuscript could therefore further benefit from these techniques and our derivation of the LD-EnKF does not preclude their application.

Sigma-Point and Unscented Versions: There have been attempts to deploy deterministic sampling for ensemble members using standard sigma-point KF [1] and unscented KF [23] approaches, with comparisons to EnKFs. We note that the deterministic sampling used in these approaches is not similar to what is proposed here. In those works, truncated (i.e., reduced order) sigma-point or unscented methods are applied; the measurement update is simply the KF update with empirical covariance and Kalman gain constructed from the quadrature based construction. The value of such approaches and whether the comparisons to EnKFs is fair has been debated [15, 24, 29].

VI. Low-Discrepancy Ensemble Kalman Filter

We now introduce the proposed LD-EnKF, that makes two important changes to the standard EnKF reviewed in Sec. IV. First, it uses deterministic samples described in Sec. VI-A instead of random samples for the ensembles. Second, it uses a more advanced form for the generalized convolution required to add noise to a given estimate. We now briefly recall necessary details on generating low-discrepancy deterministic samples and then continue to the second contribution of the new LD-EnKF.

A. Deterministic Low-Discrepancy Samples

Representing a random variable $\underline{\boldsymbol{x}}$ with pdf $f(\underline{x})$ by samples is straightforward conceptually as drawing i.i.d. random samples from a given distribution is a well-understood problem. However, the i.i.d. property is also a disadvantage as random samples do not cover the given density homogeneously, but tend to form clusters and leave holes, especially for small sample sizes. In addition, using random samples in the BRR leads to random and non-reproducible results.

A main contribution of this work is that instead of random samples for a perturbed EnKF, we will use deterministic samples generated by minimizing a distance between the given continuous density and the desired sample set. This is accomplished by means of an optimization procedure [17]. For reference, deterministic samples representing a two-dimensional standard normal distribution are shown in Fig. 2 for different sample sizes.

For optimal down-sampling (i.e., reducing the size of a given sample set), we use the procedure from [16].

B. Use of Deterministic Low-Discrepancy Samples

Instead of random samples, deterministic low-discrepancy samples as described in Sec. VI-A are used for representing all the densities occurring in the filter by ensembles: the initial state, the noises, the predicted state, and the state estimate. A deterministic representation is in line with the BRR that are fully deterministic for a given stochastic state estimation problem. Thus, random ensembles can be seen as a stochastic approximation to the deterministic densities that converge to the true solution asymptotically for a very large number of samples. This contradicts the desire for small ensemble sizes. Deterministic samples, on the other hand, naturally lead to smaller ensemble sizes as discussed in the previous section.

To fully use the enhanced representation quality of the deterministic ensembles, we propose a more advanced form of convolution compared to the standard EnKF. In the EnKF, two (generalized) convolutions of a given state distribution with a noise distribution occur in the required predictions. First, in the filter step (measurement update, assimilation step), the measurement density is predicted based on the given prior density, the measurement noise density, and the measurement model. Second, in the prediction step (time update), the state density at the next time step is calculated based on the current state density, the noise, and the process model.

These predictions can be interpreted as (generalized) convolutions of state and noise densities, with a standard convolution obtained for linear equations and additive noise. In the case of an ensemble representation of the state density, a convolution with additive noise can be viewed as "pasting" the noise density onto the location of every ensemble member leading to a mixture density. To implement a discrete convolution, the standard EnKF now only takes a single sample of each mixture component to perform the convolution (i.e., (30)). This is simple to implement, but gives poor results for small ensembles and only converges asymptotically for many samples.

Here, we propose a more advanced discrete convolution by using more samples from each mixture component or equivalently more noise samples for each ensemble member. This increases the representation quality, but also leads to an exponential increase in ensemble members. In order to keep the sample size at the desired level, we propose to use an optimal down-sampling method that reduces the ensemble size after each prediction/convolution step [16].

C. Filter Step (Measurement Update)

Similar to the standard EnKF in Sec. IV, we assume samples of the prior density $f^p(\underline{x})$ and samples of the noise density $f^v(\underline{v})$ to be given. However, we consider the more general case of different size of the prior ensemble and the noise ensemble.

Prior and noise ensembles of different sizes: The prior ensemble has L_p samples with weights

$$\underline{w}^p = \left[w_1^p, w_2^p, \cdots, w_{L_p}^p\right] \tag{31}$$

and locations

$$\mathbf{X}^p = \left[\underline{x}_1^p, \underline{x}_2^p, \cdots, \underline{x}_{L_p}^p\right], \tag{32}$$

while the noise ensemble has L_v samples with weights

$$\underline{w}^v = \left[w_1^v, w_2^v, \cdots, w_{L_v}^v\right] \tag{33}$$

and locations

$$\mathbf{V} = \left[\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_{L_v}\right]. \tag{34}$$

Temporary Posterior Ensemble with Larger Size: Similar to (27), we can now calculate a temporary posterior ensemble comprising $L_e = L_p \cdot L_v$ samples $\tilde{\underline{x}}_k^e$ by combining all noise samples with all prior samples as

$$\tilde{\underline{x}}_k^e = \underline{x}_i^p + \mathbf{K}(\underline{y} - \hat{\underline{y}}_k), \qquad \hat{\underline{y}}_k = \underline{h}(\underline{x}_i^p, \underline{v}_j), \tag{35}$$

for $i \in [1, 2, \ldots, L_p]$, $j \in [1, 2, \ldots, L_v]$, and $k = (i-1)\, L_v + j$ giving $k \in [1, 2, \ldots, L_e]$.

Down-sampling applied to the ensemble $\tilde{\mathbf{X}}^e$ gives the final ensemble $\mathbf{X}^e$ with the desired number of samples (usually L_p).

Kalman Gain: The calculation of the Kalman gain in (15) is a bit more involved in this case as we have to consider the two sample sets for prior and noise in the calculation of $\mathbf{C}^{py}$ and $\mathbf{C}^y$. The prior expected value is estimated as

$$\hat{\underline{x}}^p \approx \sum_{i=1}^{L_p} w_i^p\, \underline{x}_i^p, \tag{36}$$

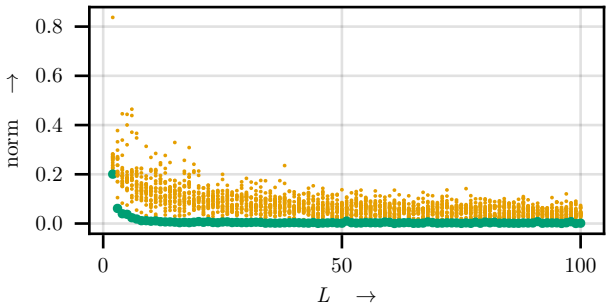

Fig. 3: Norm of covariance estimates from a single measurement update: LD-EnKF (**green**) and perturbed EnKF (**orange**) compared to the KF for ensemble sizes L ranging from 2 to 100.

and the expected value of the predicted measurements is estimated as the empirical mean based on the prior ensemble as

$$\hat{\underline{y}} \approx \sum_{i=1}^{L_p} \sum_{j=1}^{L_v} w_i^p\, w_j^v\, \underline{h}(\underline{x}_i^p, \underline{v}_j). \tag{37}$$

Based on $\hat{\underline{x}}^p$ and $\hat{\underline{y}}$, the required covariance matrices can be estimated from the ensembles as

$$\mathbf{C}^{py} \approx \sum_{i=1}^{L_p} \sum_{j=1}^{L_v} w_i^p\, w_j^v\, \left(\underline{x}_i^p - \hat{\underline{x}}^p\right)\left(\underline{h}(\underline{x}_i^p, \underline{v}_j) - \hat{\underline{y}}\right)^T \tag{38}$$

and

$$\mathbf{C}^{y} \approx \sum_{i=1}^{L_p} \sum_{j=1}^{L_v} w_i^p\, w_j^v\, \left(\underline{h}(\underline{x}_i^p, \underline{v}_j) - \hat{\underline{y}}\right)\left(\underline{h}(\underline{x}_i^p, \underline{v}_j) - \hat{\underline{y}}\right)^T. \tag{39}$$

D. Prediction Step (Time Update)

The prediction step of the LD-EnKF is analogous to the prediction step of the standard EnKF in Subsec. IV-C. The difference is that the ensemble $\mathbf{X}^e$ is of size L_e, which is typically different from the size of the process noise ensemble $\mathbf{\Xi}$ of size L_ξ. For the two ensembles $\mathbf{X}^e$ and $\mathbf{\Xi}$, the same advanced convolution is performed as described for the filter step, resulting in a temporary ensemble of size $L_e \cdot L_\xi$. This temporary ensemble is then down-sampled to the desired size, typically L_e.

VII. Evaluation

A. Single Filter Steps

We begin with a two-dimensional linear measurement equation ($N = 2$) according to (4) with measurement matrix

$$\mathbf{H} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \tag{40}$$

and noise covariance $\mathbf{R} = 0.5\ \mathrm{diag}(\underline{1}_N)$. We assume a prior with zero-mean, covariance given by an identity matrix, and a zero measurement. The true covariance matrix (according to the Kalman filter update) is then given by $\tilde{\mathbf{C}}^e = 0.2\ \mathrm{diag}(\underline{1}_N)$.

We run the proposed LD-EnKF and the perturbed EnKF for different ensemble sizes, calculate covariances of the resulting ensembles, and compare them to the true covariance matrix $\tilde{\mathbf{C}}^e$. For comparison, the Frobenius norm between estimated and true covariance matrices is calculated.

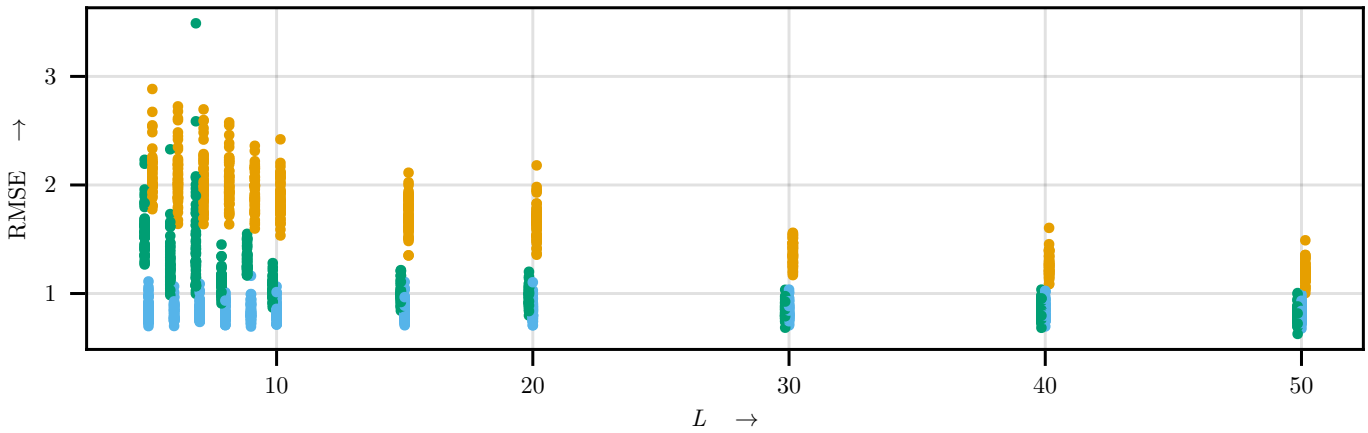

Fig. 4: Root-mean-square errors for the LD-EnKF (**green**), the perturbed EnKF with an improved sampling scheme (**orange**), and the KF (**blue**) for different ensemble sizes L.

The estimation results are shown in Fig. 3. For each ensemble size L, the LD-EnKF was run once as it produces nonrandom results (**green** dots). The perturbed EnKF was run 20 times for each L (**orange** dots). It is clear from the figure, that the LD-EnKF converges much quicker to the true covariance than the perturbed EnKF, which still shows noticeable variations even for larger ensemble sizes.

B. Heat Conductor

Inspired by [13] and [30, section V], we consider heat exchange in a one-dimensional rod of length Z. The heat distribution $T = T(t, z)$ depending on the (one-dimensional) space variable z and time t evolves according to the linear partial differential equation

$$\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial z^2} - \mu T + u(t, z), \tag{41}$$

where α is the (constant) thermal diffusivity of the medium, μ is the radiation constant, and $u(t, z)$ is a distributed external heat source/sink. We assume Neumann boundary conditions at both ends, corresponding to full insulation

$$\left.\frac{\partial T}{\partial z}\right|_{z=0} = 0 \quad \text{and} \quad \left.\frac{\partial T}{\partial z}\right|_{z=Z} = 0. \tag{42}$$

The rod is partitioned into N cells of length $\Delta z = Z/N$ with midpoints $z_i = \Delta z\,(i - 1/2)$, $i \in [1, 2, \ldots, N]$. Finite differencing in space/time yields a discrete-time process model

$$\underline{x}_{k+1} = \mathbf{A}\,\underline{x}_k + \underline{u}_k. \tag{43}$$

Here the state vector containing the discretized heat distribution is $\underline{x}_k = \left[T_{k,1}, T_{k,2}, \cdots, T_{k,N}\right]^T$, where $T_{k,i} = T(t_k, z_i)$ with $t_k = k\,\Delta t$. The system matrix is tridiagonal

$$\mathbf{A} = \begin{bmatrix} a+b & b & & & & & \\ b & a & b & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & & b & a & b & \\ & & & & b & a+b \end{bmatrix} \in \mathbb{R}^{N\times N}, \tag{44}$$

with $a = 1 - 2\,\alpha\,{}^{\Delta t}/_{(\Delta z)^2} - \mu\,\Delta t$, $b = \alpha\,{}^{\Delta t}/_{(\Delta z)^2}$, and discretized forcing term $\underline{u}_k = \left[u_{k,1}, u_{k,2}, \cdots, u_{k,L}\right]^T$ with $u_{k,i} = u(t_k, z_i)$.

For the simulation, we assume the distributed forcing term to be $u(t, Z/3) = -\cos(\pi\,t/100\,\mathrm{s})$, $u(t, 2Z/3) = \cos(\pi\,t/100\,\mathrm{s})$, and zero elsewhere. Measurements are taken at $z = Z/5$ and $z = 4Z/5$, so we have a measurement matrix $H \in \mathbb{R}^{2\times N}$.

The simulation parameters are: rod length $Z = 1\,\mathrm{m}$, time step $\Delta t = 1\,\mathrm{s}$, number of cells $N = 20$, cell width $\Delta z = Z/N$, simulation steps $n = 100$, thermal diffusivity constant $\alpha = 0.00004\,{}^{m^2}/_{s}$, and radiation constant $\mu = 0$. The entire simulation is centered around $300\,\mathrm{K}$.

Initial estimate is zero-mean Gaussian with unit covariance matrix. Measurement noise is zero-mean with $\mathbf{R} = 0.1\,\mathbf{I}_{2,2}$. We add zero-mean process noise with $Q = 0.5\,\mathbf{I}_{N,N}$. "True" trajectories for state/measurement are randomly generated.

The performance of the proposed LD-EnKF is compared to a perturbed EnKF with improved sampling described in [10]. A KF is used as a ground truth as the system is linear. The results of applying the three filters to 50 different "true" trajectories with different ensemble sizes $N \in [5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50]$ are shown in Fig. 4. The LD-EnKF clearly outperforms the perturbed EnKF with improved sampling, while the KF has the best performance, which is expected. The perturbed EnKF with purely random sampling is omitted as its performance is even worse.

VIII. Conclusions

In this manuscript, we have introduced a radically new type of EnKF. There are two main contributions. First, we employ a specific type of low-discrepancy ensemble based on optimizing a distance measure between the given continuous density (e.g., a Gaussian), and the samples. Second, we derive a more advanced implementation of the convolution of given state ensembles with the noise samples based on processing the Cartesian product of the samples with a subsequent down-sampling operation.

Compared to the typically used random ensembles, even with the improved sampling schemes in the literature, using our low-discrepancy ensembles leads to much better coverage of the state space so that far fewer samples are required for achieving the same estimation quality. In addition, the estimation results are non-random and thus reproducible.

IX. Limitations and Outlook

Due to space limitations, this paper includes a rather narrow set of simulations. Future work will see more comprehensive evaluations with more complex models including nonlinear and non-Gaussian ones. In addition, we will include comparisons against additional EnKF variants, especially the deterministic variety described in Sec. V-B.

In this paper, we used the optimization-based sampling from [17] for sampling from standard normal distributions of the noise (and the initial states). These samples are stored in a cache to avoid recalculation and transformed to the desired non-standard Gaussian. As the representation quality is slightly degraded by this transformation, an alternative is to perform on-demand optimization of sample sets for a given covariance matrix. As this is computationally more demanding, future investigations have to show if this is a worthwhile investment.

Down-sampling requires minimization of a distance measure with complexity $O(LM)$ with L, M the ensemble size before and after down-sampling. This is comparable to the complexity of the standard KF, but requires an additional computational budget beyond the standard EnKF. Future implementations of the proposed LD-EnKF will include accelerated down-sampling. The first options will use a divide-and-conquer method for the large temporary ensembles resulting during the advanced convolution step. The second option can avoid the explicit calculation of the temporary ensemble in the first place [9].

References

[1] Jaison Thomas Ambadan and Youmin Tang. "Sigma-Point Kalman Filter Data Assimilation Methods for Strongly Nonlinear Systems". *Journal of the Atmospheric Sciences* 66.2 (2009), pp. 261–285.

[2] Jeffrey L. Anderson. "An Ensemble Adjustment Kalman Filter for Data Assimilation". *Monthly Weather Review* 129.12 (2001), pp. 2884–2903.

[3] Thomas Bengtsson, Peter Bickel, and Bo Li. "Curse-of-dimensionality revisited: collapse of the particle filter in very large scale systems". English. *Probability and statistics: Essays in honor of David A. Freedman.* Beachwood, OH: IMS, Institute of Mathematical Statistics, 2008, pp. 316–315.

[4] Peter Bickel, Bo Li, and Thomas Bengtsson. "Sharp failure rates for the bootstrap particle filter in high dimensions". *Institute of Mathematical Sciences Collections* 3 (2008), pp. 318–329.

[5] Craig H. Bishop, Brian J. Etherton, and Sharanya J. Majumdar. "Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects". *Monthly Weather Review* 129.3 (Mar. 2001), pp. 420–436.

[6] Gerrit Burgers, Peter Jan van Leeuwen, and Geir Evensen. "Analysis Scheme in the Ensemble Kalman Filter". *Monthly Weather Review* 126.6 (1998), pp. 1719–1724.

[7] Mark D. Butala et al. "Asymptotic convergence of the ensemble Kalman filter". *2008 15th IEEE International Conference on Image Processing*. 2008, pp. 825–828.

[8] Philippe Courtier and Olivier Talagrand. "Variational Assimilation of Meteorological Observations With the Adjoint Vorticity Equation. II: Numerical Results". *Quarterly Journal of the Royal Meteorological Society* 113.478 (1987), pp. 1329–1347.

[9] Henning Eberhardt, Vesa Klumpp, and Uwe D. Hanebeck. "Optimal Dirac Approximation by Exploiting Independencies". *Proceedings of the 2010 American Control Conference (ACC 2010)*. Baltimore, Maryland, USA, June 2010.

[10] Geir Evensen. "Sampling strategies and square root analysis schemes for the EnKF". *Ocean Dynamics* 54.6 (2004), pp. 539–560.

[11] Geir Evensen. "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics". *Journal of Geophysical Research: Oceans* 99.C5 (1994), pp. 10143–10162.

[12] Reinhard Furrer and Thomas Bengtsson. "Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants". *Journal of Multivariate Analysis* 98.2 (2007), pp. 227–255.

[13] Humberto C. Godinez and J. David Moulton. "An Efficient Matrix-Free Algorithm for the Ensemble Kalman Filter". *Computational Geosciences* 16.3 (June 1, 2012), pp. 565–575.

[14] N. J. Gordon, D. J. Salmond, and A. F. M. Smith. "Novel approach to nonlinear/non-Gaussian Bayesian state estimation". *IEE Proceedings F - Radar and Signal Processing* 140.2 (1993), pp. 107–113.

[15] Thomas M. Hamill et al. "Comments on "Sigma-Point Kalman Filter Data Assimilation Methods for Strongly Nonlinear Systems"". *Journal of the Atmospheric Sciences* 66.11 (2009), pp. 3498–3500.

[16] Uwe D. Hanebeck. "Optimal Reduction of Multivariate Dirac Mixture Densities". *at – Automatisierungstechnik* 63.4 (Apr. 2015), pp. 265–278.

[17] Uwe D. Hanebeck, Marco F. Huber, and Vesa Klumpp. "Dirac Mixture Approximation of Multivariate Gaussian Densities". *Proceedings of the 2009 IEEE Conference on Decision and Control (CDC 2009)*. Shanghai, China, Dec. 2009.

[18] P. L. Houtekamer and Herschel L. Mitchell. "Data assimilation using an ensemble Kalman filter technique". *Monthly Weather Review* 126.3 (Mar. 1998), p. 796.

[19] P. L. Houtekamer and Fuqing Zhang. "Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation". *Monthly Weather Review* 144.12 (Dec. 2016), pp. 4489–4532.

[20] R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems". *Journal of Basic Engineering* 82.1 (Mar. 1960), pp. 35–45.

[21] W. Gregory Lawson and James A. Hansen. "Implications of Stochastic and Deterministic Filters as Ensemble-Based Data Assimilation Methods in Varying Regimes of Error Growth". *Monthly Weather Review* 132.8 (2004), pp. 1966–1981.

[22] François Le Gland, Valérie Monbet, and Vu-Duc Tran. *Large sample asymptotics for the ensemble Kalman filter*. Research Report RR-7014. INRIA, 2009, p. 25.

[23] X. Luo and I.M. Moroz. "Ensemble Kalman filter with the unscented transform". *Physica D: Nonlinear Phenomena* 238.5 (2009), pp. 549–562.

[24] X. Luo, I.M. Moroz, and I. Hoteit. "Reply to "Comment on 'Ensemble Kalman filter with the unscented transform'"". *Physica D: Nonlinear Phenomena* 239.17 (2010), pp. 1662–1664.

[25] Jan Mandel, Loren Cobb, and Jonathan D. Beezley. "On the convergence of the ensemble Kalman filter". *Applications of Mathematics* 56.6 (2011), pp. 533–541.

[26] Lars Nerger et al. "A Unification of Ensemble Square Root Kalman Filters". *Monthly Weather Review* 140.7 (2012), pp. 2335–2345.

[27] Dinh Tuan Pham. "Stochastic Methods for Sequential Data Assimilation in Strongly Nonlinear Systems". *Monthly Weather Review* 129.5 (2001), pp. 1194–1207.

[28] Michael Roth et al. "The Ensemble Kalman filter: a signal processing perspective". *EURASIP Journal on Advances in Signal Processing* 2017.1 (2017), p. 56.

[29] Pavel Sakov. "Comment on "Ensemble Kalman filter with the unscented transform"". *Physica D: Nonlinear Phenomena* 238.22 (2009), pp. 2227–2228.

[30] Joris Sijs, Uwe D. Hanebeck, and Benjamin Noack. "An Empirical Method to Fuse Partially Overlapping State Vectors for Distributed State Estimation". *Proceedings of the 2013 European Control Conference (ECC 2013)*. Zürich, Switzerland, July 2013.

[31] Chris Snyder et al. "Obstacles to High-Dimensional Particle Filtering". *Monthly Weather Review* 136.12 (2008), pp. 4629–4640.

[32] Olivier Talagrand and Philippe Courtier. "Variational Assimilation of Meteorological Observations With the Adjoint Vorticity Equation. I: Theory". *Quarterly Journal of the Royal Meteorological Society* 113.478 (1987), pp. 1311–1328.

[33] Michael K. Tippett et al. "Ensemble Square Root Filters". *Monthly Weather Review* 131.7 (2003), pp. 1485–1490.

[34] Peter Jan van Leeuwen. "A consistent interpretation of the stochastic version of the Ensemble Kalman Filter". *Quarterly Journal of the Royal Meteorological Society* 146.731 (2020), pp. 2815–2825.

[35] Sanita Vetra-Carvalho et al. "State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems". *Tellus B: Chemical and Physical Meteorology* (Jan. 2018).

[36] Jeffrey S. Whitaker and Thomas M. Hamill. "Ensemble Data Assimilation without Perturbed Observations". *Monthly Weather Review* 130.7 (2002), pp. 1913–1924.