THE ENSEMBLE KALMAN FILTER AND ITS RELATIONS TO OTHER NONLINEAR FILTERS

Michael Roth, Carsten Fritsche*, Gustaf Hendeby*†, and Fredrik Gustafsson**

* Linköping University, Department of Electrical Engineering, Linköping, Sweden,
e-mail: {roth,carsten,hendeby,fredrik}@isy.liu.se
† Swedish Defence Research Agency (FOI), Linköping, Sweden
e-mail: gustaf.hendeby@foi.se

ABSTRACT

The Ensemble Kalman filter (EnKF) is a standard algorithm in oceanography and meteorology, where it has got thousands of citations. It is in these communities appreciated since it scales much better with state dimension n than the standard Kalman filter (KF). In short, the EnKF propagates ensembles with N state realizations instead of mean values and covariance matrices and thereby avoids the computational and storage burden of working on $n \times n$ matrices. Perhaps surprising, very little attention has been devoted to the EnKF in the signal processing community. In an attempt to change this, we present the EnKF in a Kalman filtering context. Furthermore, its application to nonlinear problems is compared to sigma point Kalman filters and the particle filter, so as to reveal new insights and improvements for high-dimensional filtering algorithms in general. A simulation example shows the EnKF performance in a space debris tracking application.

***Index Terms*—** Kalman filter, ensemble Kalman filter, sigma point Kalman filter, UKF, particle filter

1. INTRODUCTION

The Ensemble Kalman Filter (EnKF) is an algorithm for state estimation in high-dimensional state-space models. Its development has been driven by applications in oceanography, meteorology, and other geosciences, in which the state dimension can be in the order of millions [1].

With the ever increasing amount of available data, such high-dimensional state estimation problems also become more important in the signal processing community. Still, the EnKF has been widely overlooked by signal processing researchers. It is our intention to change this, first, by giving a clear presentation that derives the EnKF from the linear Kalman filter (KF) and, second, by giving some relations to more familiar algorithms such as sigma point Kalman filters (*e.g.*, the UKF) and the particle filter (PF). We hope that this leads to a cross-fertilization of ideas that will reveal new insights and improvements, both for the EnKF and filtering algorithms for high-dimensional problems in general.

One reason for the lack of awareness towards the EnKF is that its development took place in geoscientific rather than signal processing journals. For example, the first EnKF paper [2] was published in a geophysics journal; the paper [3] which contains an important correction to [2] appeared in a meteorological journal. Accordingly, application specific jargon is often used: The KF time update or prediction step is often termed forecast in EnKF literature; the measurement update or correction step is termed analysis or data assimilation. An extensive listing of geoscientific EnKF publications is given in [4]. Popular KF references such as [5] or even Kalman's paper [6] are seldom cited in EnKF articles.

Some attention beyond the geoscientific context has been devoted to the EnKF by the automatic control [7, 8] and statistics communities [9–12]. In particular, a result that for linear systems the EnKF converges to the KF as the number of ensemble members tends to infinity has been reported by different authors [12–14] rather recently.

The structure of the paper is as follows. Sec. 2 develops the EnKF from the KF for linear systems. Sec. 3 shows how the EnKF can be applied to nonlinear systems, and establishes relations to sigma point Kalman filters and the particle filter. A high-dimensional simulation example in which space debris is tracked is presented in Sec. 4, and followed by some concluding remarks.

2. FROM KALMAN FILTER TO ENSEMBLE KALMAN FILTER IN LINEAR SYSTEMS

This section recalls the KF which is then used to derive the ensemble updates in the EnKF. Furthermore, we discuss the filter gain computation and several practical aspects.

This work was supported by the project Scalable Kalman Filters granted by the Swedish Research Council.

2.1. The Kalman Filter

We consider linear stochastic state-space models

$$x_{k+1} = Fx_k + Gv_k, \tag{1a}$$
$$y_k = Hx_k + e_k, \tag{1b}$$

with an n-dimensional state x, an m-dimensional measurement y, with $\mathrm{E}(x_0) = \hat{x}_0$, $\mathrm{cov}(x_0) = P_0$, $\mathrm{cov}(v_k) = Q$ and $\mathrm{cov}(e_k) = R$. The Kalman filter [5] for (1) can be divided into a time update of the state estimate and its covariance

$$\hat{x}_{k+1|k} = F\hat{x}_{k|k}, \tag{2a}$$
$$P_{k+1|k} = FP_{k|k}F^T + GQG^T; \tag{2b}$$

the computation of the predicted output and its covariance

$$\hat{y}_{k|k-1} = H\hat{x}_{k|k-1}, \tag{3a}$$
$$S_k = HP_{k|k-1}H^T + R; \tag{3b}$$

and a measurement update of the state estimate and its covariance

$$K_k = P_{k|k-1}H^T S_k^{-1} = M_k S_k^{-1}, \tag{4a}$$
$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1}) \tag{4b}$$
$$= (I - K_kH)\hat{x}_{k|k-1} + K_ky_k, \tag{4c}$$
$$P_{k|k} = (I - K_kH)P_{k|k-1}(I - K_kH)^T + K_kRK_k^T. \tag{4d}$$

Here, $P_{k|k}$ is written in Joseph form which appears useful in the upcoming discussion of the EnKF. It should be noted that the inverse S_k^{-1} in the gain computation (4a) does not need to be computed explicitly. Rather, it is numerically advisable to compute K_k by solving the linear system of equations $K_kS_k = M_k$ [15].

2.2. The Ensemble Kalman Filter

The KF is problematic in high-dimensional state-spaces because it requires storing and processing $n \times n$ covariance matrices. Moreover, the full $P_{k|k}$ and $P_{k|k-1}$ are hardly ever useful as output information to the user. These motives led to the development of the EnKF, which is based on the idea to condense the information that is carried by $\hat{x}_{k|k}$ and $P_{k|k}$ into samples. Specifically, the EnKF represents the filtering result by an ensemble of N realizations $\{x_k^{(i)}\}_{i=1}^N$ such that

$$\bar{x}_{k|k} = \tfrac{1}{N}\sum_{i=1}^{N} x_k^{(i)} \approx \hat{x}_{k|k}, \tag{5a}$$
$$\bar{P}_{k|k} = \tfrac{1}{N-1}\sum_{i=1}^{N}\big(x_k^{(i)} - \bar{x}_{k|k}\big)\big(x_k^{(i)} - \bar{x}_{k|k}\big)^T \approx P_{k|k}. \tag{5b}$$

The ensemble can be stored in an $n \times N$ matrix $X_{k|k}$, which also allows for the compact notation of the ensemble mean and (unbiased) ensemble covariance

$$\bar{x}_{k|k} = \tfrac{1}{N}X_{k|k}\mathbb{1}, \tag{6a}$$
$$\bar{P}_{k|k} = \tfrac{1}{N-1}\tilde{X}_{k|k}\tilde{X}_{k|k}^T, \tag{6b}$$

where $\mathbb{1} = [1, \ldots, 1]^T$ is an N-dimensional column vector and $\tilde{X}_{k|k} = X_{k|k} - \bar{x}_{k|k}\mathbb{1}^T$ is an ensemble of deviations from $\bar{x}_{k|k}$. It turns out that computation and storage of the large matrices in (5b) or (6b) is avoided in the EnKF, which makes the algorithm attractive for solving high-dimensional state estimation problems.

2.3. Ensemble propagation for a known Kalman gain

The EnKF time update corresponds to (2) and amounts to computing an ensemble $X_{k+1|k}$ that encodes $\hat{x}_{k+1|k}$ and $P_{k+1|k}$ from $X_{k|k}$. This can be achieved with an ensemble $V_k = [v_k^{(1)}, \ldots, v_k^{(N)}]$ of N process noise realizations:

$$X_{k+1|k} = FX_{k|k} + GV_k. \tag{7}$$

The EnKF time update amounts to a simulation. Consequently, also continuous time and/or nonlinear dynamics can be considered as long as state transitions can be simulated. In fact, the time update is typically omitted in the EnKF literature [8], where an ensemble $X_{k|k-1}$ is often the starting point.

In the EnKF measurement update, a prediction ensemble $X_{k|k-1}$ is processed to obtain the filtering ensemble $X_{k|k}$ that encodes the KF mean and covariance. For the moment, we assume that the Kalman gain K_k is known but discuss its computation in Sec. 2.4.

With K_k available, the KF update (4c) can be applied to each ensemble member as follows

$$X_{k|k} = (I - K_kH)X_{k|k-1} + K_ky_k\mathbb{1}^T. \tag{8}$$

The resulting ensemble average (6a) is the correct $\hat{x}_{k|k}$. However, with $y_k\mathbb{1}^T$ known, the sample covariance of $X_{k|k}$ gives only the first term of (4d) and therefore fails to encode $P_{k|k}$. A solution [3] is to account for the missing term $K_kRK_k^T$ by adding artificial zero-mean measurement noise realizations $E_k = [e_k^{(1)}, \ldots, e_k^{(N)}]$ with covariance R, according to

$$X_{k|k} = (I - K_kH)X_{k|k-1} + K_ky_k\mathbb{1}^T - K_kE_k. \tag{9}$$

Then, $X_{k|k}$ correctly encodes $\hat{x}_{k|k}$ and $P_{k|k}$. The model (1) is implicit in (9) because the matrix H appears. If we, similar to (3), define a predicted output ensemble

$$Y_{k|k-1} = HX_{k|k-1} + E_k \tag{10}$$

that encodes $\hat{y}_{k|k-1}$ and S_k, we can reformulate (9) to an update that resembles (4b):

$$X_{k|k} = X_{k|k-1} + K_k(y_k\mathbb{1}^T - Y_{k|k-1}). \tag{11}$$

In contrast to the update (9), (11) is merely sample based, which is a useful property for the extension to nonlinear systems.

2.4. Estimating the Kalman Gain

The preceding derivation of the EnKF assumed a known gain K_k, which in the KF is a function of covariance matrices. Accordingly, a more explicit version of (4a) can be written as

$$K_k = M_k S_k^{-1} = \operatorname{cov}(x_k, y_k \mid y_{1:k-1}) \operatorname{cov}(y_k \mid y_{1:k-1})^{-1}. \tag{12}$$

The EnKF encodes predicted mean values and covariance matrices via ensembles of random states (7) and outputs (10). The fully deterministic K_k of (4a) is in (9) or (11) replaced by the EnKF gain $\bar{K}_k = \bar{M}_k \bar{S}_k^{-1}$. The randomness in $\bar{M}_k$ and $\bar{S}_k$ implies a random $\bar{K}_k$. Below, we discuss methods for computing the matrices $\bar{M}_k$ and $\bar{S}_k$.

A direct way to compute $\bar{M}_k$ and $\bar{S}_k$ is to utilize $X_{k|k-1}$ from (7) and $Y_{k|k-1}$ from (10). First, the ensemble averages $\bar{x}_{k|k-1} = \frac{1}{N} X_{k|k-1} \mathbb{1}$ and $\bar{y}_{k|k-1} = \frac{1}{N} Y_{k|k-1} \mathbb{1}$ are used to compute the deviation ensembles

$$\tilde{X}_{k|k-1} = X_{k|k-1} - \bar{x}_{k|k-1} \mathbb{1}^T, \tag{13a}$$

$$\tilde{Y}_{k|k-1} = Y_{k|k-1} - \bar{y}_{k|k-1} \mathbb{1}^T, \tag{13b}$$

from which we can then obtain

$$\bar{M}_k = \tfrac{1}{N-1} \tilde{X}_{k|k-1} \tilde{Y}_{k|k-1}^T, \tag{13c}$$

$$\bar{S}_k = \tfrac{1}{N-1} \tilde{Y}_{k|k-1} \tilde{Y}_{k|k-1}^T. \tag{13d}$$

The computations (13) are entirely sampling based, which is useful for the upcoming nonlinear case. However, they inherit errors due to the random sampling that can be avoided.

The basis for the KF is the knowledge how mean values and covariance matrices propagate through linear systems (1). Consequently, we can use model knowledge to decrease the randomness in (13). First, we can replace the ensemble averages in (13a) and (13b) by $\bar{x}_{k|k-1} = F \bar{x}_{k-1|k-1}$ and $\bar{y}_{k|k-1} = HF \bar{x}_{k-1|k-1}$, with a previous filtering mean $\bar{x}_{k-1|k-1}$. Second, we know that the measurement noise e_k is not correlated with the state and should not influence $\bar{M}_k$. Therefore, we can work with an unperturbed output deviation ensemble

$$\tilde{Z}_{k|k-1} = H X_{k|k-1} - \bar{y}_{k|k-1} \mathbb{1}^T \tag{14a}$$

to compute

$$\bar{M}_k = \tfrac{1}{N-1} \tilde{X}_{k|k-1} \tilde{Z}_{k|k-1}^T, \tag{14b}$$

$$\bar{S}_k = \tfrac{1}{N-1} \tilde{Z}_{k|k-1} \tilde{Z}_{k|k-1}^T + R. \tag{14c}$$

It should be noted that the above considerations do not apply to the nonlinear systems of the next section. Neither (13) nor (14) require the computation of $n \times n$ matrices.

2.5. Practical aspects

The practical challenges and remedies below have their origin in geoscientific applications, where the state-space models often come from discretizing partial differential equations for spatially distributed grid points. They are described in a dedicated chapter in [4].

The matrices $\bar{M}_k$ and $\bar{S}_k$ are estimated from data and subject to random sampling errors. For example, there might be non-zero elements in $\bar{M}_k$ where M_k is actually zero. This can lead to unwanted correlation among state components, an effect called spurious correlation in EnKF literature. A suggested remedy is called localization and is based on the fact that most measurements only affect a small subset of the entire state vector. The idea is to carry out the measurement update only for these specific states.

Furthermore, heuristics that rely on artificially increasing the uncertainty are used to improve filter performance. The method is called covariance inflation in EnKF literature and amounts to increasing the spread of the ensembles. A similarity can be seen to heuristics in the filtering literature, such as the "fudge factor" of [16] and dithering in the PF [17].

3. THE ENSEMBLE KALMAN FILTER FOR NONLINEAR SYSTEMS

The filtering problem in nonlinear and/or non-Gaussian state-space models

$$x_{k+1} = f(x_k, v_k), \tag{15a}$$

$$y_k = h(x_k, e_k), \tag{15b}$$

with known probability density functions $p(x_0)$, $p(v_k)$, and $p(e_k)$ has an elegant conceptual solution: the Bayesian filtering equations [18] recursively yield the posterior $p(x_k \mid y_{1:k})$. Unfortunately, the required computations are intractable in all but the simplest cases.

We have shown in the previous section that the EnKF can be executed entirely sampling based. By overloading the functions in (15) to accept ensembles as input, we see from

$$X_{k+1|k} = f(X_{k|k}, V_k), \tag{16a}$$

$$Y_{k|k-1} = h(X_{k|k-1}, E_k) \tag{16b}$$

that the EnKF with the measurement update (11) is applicable to nonlinear systems.

This section highlights parallels between the EnKF and two other (approximate) nonlinear filtering algorithms, sigma point Kalman filters and the particle filter.

3.1. The EnKF and sigma point Kalman filters

Another class of sampling based algorithms that can be applied to (15) are sigma point Kalman filters such as the unscented KF [19, 20], or variants that are based on interpolation [21] or numerical integration [21, 22]. All of [19–22] share the measurement update (4), that yields an estimate $\hat{x}_{k|k}$ and a covariance $P_{k|k}$, with the KF. Hence, sigma point Kalman filters share the storage requirements of the KF, and inherit the problems that appear for large n.

A similarity between EnKF and sigma point filters is that samples are propagated according to (16). These samples are called ensemble in the former and sigma points in the latter class of filters. In contrast to the EnKF that retains the ensemble, the sigma points are generated from $\hat{x}_{k|k}$ and $P_{k|k}$ at the beginning of each iteration and also condensed into mean values and covariance matrices before the measurement update. Furthermore, each sigma point is assigned a weight. In contrast to the random sampling in the EnKF, the sigma point and weight generation is fully deterministic in all of [19–22]. Based on the viewpoint in the algorithm development, however, the sampling methodology and sample size differs.

As illustration, we sketch one possible UKF time update for (15). Let r be the dimension of $[x_k^T, v_k^T]^T$ and $N = 2r+1$. First, N sigma points $X_{k|k}$ and V_k, and an N-dimensional weight vector $w_{k|k}$ are generated such that $X_{k|k}w_{k|k} = \hat{x}_{k|k}$, $V_k w_{k|k} = 0$, and with $W_{k|k} = \operatorname{diag}(w_{k|k})$

$$\begin{bmatrix} X_{k|k} - \hat{x}_{k|k}\mathbb{1}^T \\ V_k \end{bmatrix} W_k \begin{bmatrix} X_{k|k} - \hat{x}_{k|k}\mathbb{1}^T \\ V_k \end{bmatrix}^T = \begin{bmatrix} P_{k|k} & 0 \\ 0 & Q \end{bmatrix}. \quad (17)$$

Second, the transformed sigma points $X_{k+1|k}$ are generated via (16a). In contrast to the EnKF, however, $X_{k+1|k}$ is condensed into a mean value $\hat{x}_{k+1|k} = X_{k+1|k}w_{k|k}$ and $P_{k+1|k} = X_{k+1|k}(W_{k|k} - w_{k|k}w_{k|k}^T)X_{k+1|k}$. The latter expression compactly denotes a weighted sample covariance.

A similar procedure yields M_k and S_k, and eventually the Kalman gain K_k. There is no randomness in the algorithm and for linear systems, the UKF is equivalent to the KF. The UKF gain computation is in fact similar to the one in EnKF, except that all ensemble members are equally weighted in the latter.

3.2. The EnKF and the Particle Filter

Similar to the EnKF, the particle filter (PF) retains N samples, called particles, throughout each iteration. The algorithm has been developed to approximately solve the Bayesian filtering equations [17, 23]. Using the principle of importance sampling, the PF generates samples $X_{k|k-1}$ from a proposal distribution and then assigns weights $w_{k|k}$ such that weighted particle averages approximate expected values with respect to the posterior $p(x_k \mid y_{1:k})$.

The key difference between the EnKF and the PF is that the PF samples are only generated once in each iteration, and then left unchanged. In basic PF variants, the measurement influences merely the weights $w_{k|k}$, and the particles only indirectly through a resampling step that is carried out to avoid the problem of weight degeneracy [17]. More advanced PF variants can make use of y_k in the generation of samples [23]. The EnKF, in contrast, changes each ensemble member in the measurement update.

For a specific choice of PF proposal distribution, characterized by the transition density $p(x_{k+1} \mid x_k)$, the PF time update amounts to a simulation of each particle with an independent process noise realization similar to (16a). That is, the EnKF and PF time updates are similar.

Particle filters are capable of solving severely nonlinear and non-Gaussian problems, but can only be applied to low-dimensional problems without further care. One reason for this is the problem of generating meaningful samples in high-dimensional spaces.

4. ENKF PERFORMANCE IN SIMULATIONS

Here a simulation is presented inspired by the problem to track space debris, a growing problem in space aviation. The example has been considerably simplified not to draw attention from the EnKF itself. Hence, the objects travel in 1D and they are affected by the same external force, the solar wind, effectively changing their speed. The nominal angular velocity of the object is chosen to represent debris orbiting earth at a constant altitude of 100 km.

The trajectory of object i is modeled with a constant velocity model, $x^i = [\theta^i, \omega^i]^T$,

$$x_{k+1}^i = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x_k^i + \begin{bmatrix} 0 \\ T \end{bmatrix} v_k, \quad (18a)$$

$$y_k^i = \begin{bmatrix} 1 & 0 \end{bmatrix} x_k^i + e_k^i. \quad (18b)$$

The sample interval is $T = 1$ min, and measurements are available every 10 min. At these times all objects can be observed. The nominal object speed is $\omega^n = 1.2 \cdot 10^{-5}$ rad/s, the measurement noise is $R^i = \operatorname{cov}(e_k^i) = (1^\circ)^2$, the common process noise is $Q = \operatorname{cov}(v_k) = (10^{-4}/60)^2$, and the initial uncertainty of each object state is $P_0^i = \operatorname{cov}(x_0^i) = 10^6 \operatorname{diag}(R^i, Q)$.

The individual states are stacked to handle M objects. The resulting model can be expressed using Kronecker product notation

$$x_{k+1} = I_M \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x_k + \mathbb{1}_M \otimes \begin{bmatrix} 0 \\ T \end{bmatrix} v_k, \quad (19a)$$

$$y_k = I_M \otimes \begin{bmatrix} 1 & 0 \end{bmatrix} x_k + e_k, \quad (19b)$$

where x has the dimension $n = 2M$.

100 Monte Carlo simulations each were performed for different numbers of tracked objects M and different ensemble sizes N. The result is presented in Fig. 1. The illustrated MSE is computed as the average squared position error of all objects over time. Of the total simulation time of 1 000 min only the last 250 min are used for evaluation.

The result verifies that the EnKF converges to the KF as N grows. However, as M and thus n increase, the ensemble size N has to grow significantly in order to retain the KF performance for the chosen example.

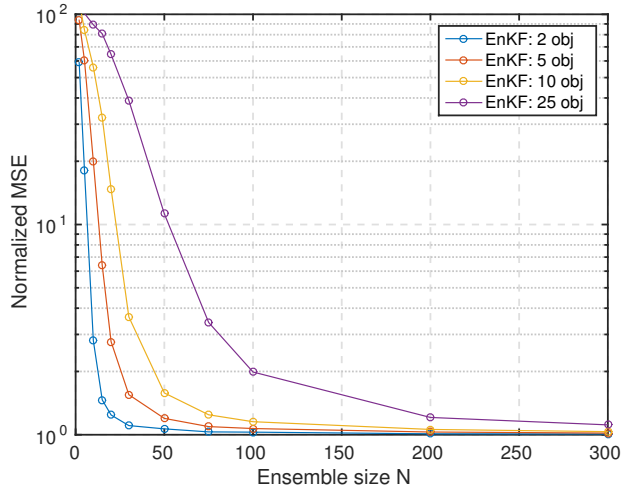

Fig. 1. Time average of the position MSE as a function of the ensemble size N, normalized by the KF MSE, for different numbers of objects M. The state dimension is $n = 2M$. The results are based on 100 Monte Carlo simulations.

5. CONCLUDING REMARKS

We have presented the ensemble Kalman filter (EnKF) in a way that clearly shows its origin in the Kalman filter (KF), and makes it easily accessible to the signal processing community. We have discussed how the EnKF can be applied to nonlinear non-Gaussian filtering problems and at the same time highlighted the similarities and differences to sigma point Kalman filters and the particle filter (PF). The simulation experiment shows that the EnKF gives the KF results for linear problems, once the ensemble becomes large enough.

The EnKF is still an unexplored method in the signal processing community. We have provided some initial insights into the algorithm. However, we believe that this is only the beginning and that cross-fertilization with other methods is possible. Understanding the Kalman gain approximation could be one key ingredient in this. Moreover, it is important to characterize the model features that make the EnKF work well in its successful applications.

References

[1] J. L. Anderson, "Ensemble Kalman filters for large geophysical applications," *IEEE Control Syst. Mag.*, vol. 29, no. 3, pp. 66–82, Jun. 2009.

[2] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research*, vol. 99, no. C5, pp. 10 143–10 162, 1994.

[3] G. Burgers, P. J. van Leeuwen, and G. Evensen, "Analysis scheme in the ensemble Kalman filter," *Monthly Weather Review*, vol. 126, no. 6, pp. 1719–1724, Jun. 1998.

[4] G. Evensen, *Data Assimilation: The Ensemble Kalman Filter*, 2nd ed. Dordrecht; New York: Springer, Aug. 2009.

[5] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice Hall, Jun. 1979.

[6] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, Mar. 1960.

[7] S. Gillijns, O. B. Mendoza, J. Chandrasekar, B. L. R. De Moor, D. S. Bernstein, and A. Ridley, "What is the ensemble Kalman filter and how well does it work?" in *Proceedings of the American Control Conference*, Jun. 2006.

[8] G. Evensen, "The ensemble Kalman filter for combined state and parameter estimation," *IEEE Control Syst. Mag.*, vol. 29, no. 3, pp. 83–104, Jun. 2009.

[9] R. Furrer and T. Bengtsson, "Estimation of high dimensional prior and posterior covariance matrices in Kalman filter variants," *Journal of Multivariate Analysis*, vol. 98, no. 6, pp. 227–255, Feb. 2007.

[10] M. Frei and H. R. Künsch, "Mixture ensemble Kalman filters," *Computational Statistics & Data Analysis*, vol. 58, pp. 127–138, Feb. 2013.

[11] M. Frei, "Ensemble Kalman filtering and generalizations," Dissertation, ETH, Zürich, 2013, nr. 21266.

[12] F. Le Gland, V. Monbet, and V. Tran, "Large sample asymptotics for the ensemble Kalman filter," in *The Oxford Handbook of Nonlinear Filtering*, D. Crisan and B. Rozovskii, Eds. Oxford University Press, 2011, ch. 7.3, pp. 598–634.

[13] M. D. Butala, J. Yun, Y. Chen, R. A. Frazin, and F. Kamalabadi, "Asymptotic convergence of the ensemble Kalman filter," in *15th IEEE International Conference on Image Processing*, Oct. 2008, pp. 825–828.

[14] J. Mandel, L. Cobb, and J. D. Beezley, "On the convergence of the ensemble Kalman filter," *Applications of Mathematics*, vol. 56, no. 6, pp. 533–541, Jan. 2011.

[15] L. N. Trefethen and D. Bau, III, *Numerical Linear Algebra*. Philadelphia: SIAM, Jun. 1997.

[16] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. Wiley-Interscience, Jun. 2001.

[17] F. Gustafsson, "Particle filter theory and practice with positioning applications," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 25, no. 7, pp. 53–82, 2010.

[18] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, Mar. 1970.

[19] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proceedings of the American Control Conference*, vol. 3, 1995, pp. 1628–1632.

[20] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401– 422, Mar. 2004.

[21] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Autom. Control*, vol. 45, no. 5, pp. 910–927, May 2000.

[22] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.

[23] A. Doucet, S. J. Godsill, and C. Andrieu, "On sequential Monte Carlo methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.