



Particle Kalman Filtering: A Nonlinear Framework for Ensemble Kalman Filters

Ibrahim Hoteit, Xiaodong Luo, Dinh-Tuan Pham, and Irene M. Moroz

Citation: [AIP Conference Proceedings](#) **1281**, 1075 (2010); doi: 10.1063/1.3497823

View online: <http://dx.doi.org/10.1063/1.3497823>

View Table of Contents: <http://scitation.aip.org/content/aip/proceeding/aipcp/1281?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

[Tracking of time-evolving sound speed profiles in shallow water using an ensemble Kalman-particle filter](#)
J. Acoust. Soc. Am. **133**, 1377 (2013); 10.1121/1.4790354

[Nonlinear Continuous Global Optimization by Modified Differential Evolution](#)
AIP Conf. Proc. **1281**, 955 (2010); 10.1063/1.3498653

[Ensemble Kalman and \$H_\infty\$ Filters](#)
AIP Conf. Proc. **1281**, 19 (2010); 10.1063/1.3498332

[On Ensemble Nonlinear Kalman Filtering with Symmetric Analysis Ensembles](#)
AIP Conf. Proc. **1281**, 1088 (2010); 10.1063/1.3497831

[Tracking of geoacoustic parameters using Kalman and particle filters](#)
J. Acoust. Soc. Am. **125**, 746 (2009); 10.1121/1.3050280

Particle Kalman Filtering: A Nonlinear Framework for Ensemble Kalman Filters

Ibrahim Hoteit*, Xiaodong Luo^{*,†}, Dinh-Tuan Pham^{**} and Irene M. Moroz[‡]

^{*}King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

[†]The Oxford-Man Institute, Eagle House, Walton Well Road, Oxford, 6ED, UK

^{**}Centre National de la Recherche Scientifique (CNRS), Grenoble, France

[‡]Mathematical Institute, 24-29 St Giles', Oxford, OX1 3LB, UK

Abstract. Optimal nonlinear filtering consists of sequentially determining the conditional probability distribution functions (pdf) of the system state, given the information of the dynamical and measurement processes and the previous measurements. Once the pdfs are obtained, one can determine different estimates, for instance, the minimum variance estimate, or the maximum a posteriori estimate, of the system state. It can be shown that, many filters, including the Kalman filter (KF) and the particle filter (PF), can be derived based on this sequential Bayesian estimation framework.

In this contribution, we present a Gaussian mixture-based framework, called the particle Kalman filter (PKF), and discuss how the different EnKF methods can be derived as simplified variants of the PKF. We also discuss approaches to reducing the computational burden of the PKF in order to make it suitable for complex geosciences applications. We use the strongly nonlinear Lorenz-96 model to illustrate the performance of the PKF.

Keywords: Data assimilation; Ensemble Kalman filter; Particle Kalman filter

PACS: 92.60Wc, 02.50-r

INTRODUCTION

Data assimilation in geophysical systems often encounters the following problems: (a) the system in assimilation is nonlinear (nonlinearity); and (b) the system in assimilation is high-dimensional (high-dimensionality). Nonlinearity often implies that it is difficult to obtain an exact solution of the data assimilation problem. Instead, one has to be contented with some approximate solution and tolerates the resulting approximation error. On the other hand, high-dimensionality often makes it computationally intensive to execute an assimilation scheme. One may have to make further simplifications for speeding-up.

The ensemble Kalman filter (EnKF) [1, 2, 3, 4] is a data assimilation method which attempts to tackle some of the above problems. Essentially, the EnKF is a Monte Carlo implementation of the Kalman filter (KF), with an ensemble of the system state as the representation of the whole state space. With a typically small ensemble size, the computational cost of the EnKF is reasonably cheap. Moreover, by propagating an ensemble of the system states forward through the governing equations of a dynamical system and evaluating the statistics (e.g. sample mean and covariance) of the system states based on the propagated ensemble, the EnKF can also “bypass” the problem of nonlinearity in the sense that it does not require to linearize the nonlinear system as does the extended Kalman filter. However, the problem of non-Gaussianity is not fully addressed in the EnKF. Instead, it is customary to assume that, both the dynamical and observation noise, and the system state (approximately) follow some Gaussian distributions. In practice, this assumption may not always be true. However, due to its simplicity in implementation and the ability to achieve reasonable accuracy with relatively low computational cost in many situations, the EnKF is one of the most popular methods for data assimilation in high-dimensional systems.

In this work we apply a framework, called the particle Kalman filter (PKF) [5], for data assimilation in nonlinear/non-Gaussian systems. The main idea of the PKF is to approximate the pdf of the system state by a set of Gaussian distributions, each of them corresponding to an EnKF. Moreover, the EnKFs in the PKF can be implemented in a parallel way, so that the computational speed of the PKF may be comparable to that of an individual EnKF if the computational resources are available.

The organization of this work is as follows. We first present the data assimilation problem in consideration and give a conceptual solution based on the framework of recursive Bayesian estimation (RBE). We then consider a (practical) approximate solution, the PKF, to the data assimilation problem. Finally we use a numerical example to illustrate the

performance of the PKF.

PROBLEM STATEMENT AND THE CONCEPTUAL SOLUTION

We consider the data assimilation problem in the following scenario:

$$\mathbf{x}_k = \mathcal{M}_{k,k-1}(\mathbf{x}_{k-1}) + \mathbf{u}_k, \quad (1a)$$

$$\mathbf{y}_k = \mathcal{H}_k(\mathbf{x}_k) + \mathbf{v}_k, \quad (1b)$$

where the transition operator $\mathcal{M}_{k,k-1}$ and the observation operator \mathcal{H}_k are both possibly nonlinear. The dynamical and observation noise, in terms of \mathbf{u}_k and \mathbf{v}_k respectively, are Gaussian, and their pdfs, $p(\mathbf{u}_k)$ and $p(\mathbf{v}_k)$, are assumed to be known to us. One may also consider the scenario that \mathbf{u}_k and/or \mathbf{v}_k are (is) non-Gaussian, see, for example, [6].

The problem of our interest is to estimate the system state \mathbf{x}_k at time k , given the historical observations $\mathbf{Y}_k = \{\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k\}$ up to and including time k , and the prior pdf $p(\mathbf{x}_i | \mathbf{Y}_{i-1})$ of the system state \mathbf{x}_i at some instant i ($i \leq k$).

Recursive Bayesian estimation [7] provides a framework that recursively solves the above problem in terms of some conditional pdfs. Let $p(\mathbf{x}_k | \mathbf{Y}_{k-1})$ be the prior pdf of the state \mathbf{x}_k conditioned on the observations \mathbf{Y}_{k-1} . Once the new observation \mathbf{y}_k is available, one updates the prior pdf to the posterior $p(\mathbf{x}_k | \mathbf{Y}_k)$ according to Bayes' rule. Then by evolving the state \mathbf{x}_k forward through the system model Eq. (1a), one computes the prior pdf $p(\mathbf{x}_{k+1} | \mathbf{Y}_k)$ at the next time instant. Concretely, one may formulate the mathematical description of the aforementioned idea as follows:

$$p(\mathbf{x}_k | \mathbf{Y}_{k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{Y}_{k-1}) d\mathbf{x}_{k-1}, \quad (2a)$$

$$p(\mathbf{x}_k | \mathbf{Y}_k) = \frac{p(\mathbf{y}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{Y}_{k-1})}{\int p(\mathbf{y}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{Y}_{k-1}) d\mathbf{x}_k}, \quad (2b)$$

where $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ is equal to the value of $p(\mathbf{u}_k)$ evaluated at $\mathbf{u}_k = \mathbf{x}_k - \mathcal{M}_{k-1,k}(\mathbf{x}_{k-1})$ (by Eq. (1a)) and conditioned on \mathbf{x}_{k-1} , and $p(\mathbf{y}_k | \mathbf{x}_k)$ is equal to the value of $p(\mathbf{v}_k)$ evaluated at $\mathbf{v}_k = \mathbf{y}_k - \mathcal{H}_k(\mathbf{x}_k)$ (by Eq. (1b)) and conditioned on \mathbf{x}_k . Once the conditional pdfs in Eq. (2) have been obtained, all of the statistical information of interest, for example, the conditional means, can be evaluated based on the explicit forms of the pdfs.

Note that Eq. (2) only provides a conceptual framework for pdf estimations. In many situations, the integrals in Eq. (2) are intractable. Thus one may have to resort to some approximation method to solve Eq. (2), as will be shown later.

PARTICLE KALMAN FILTER AS THE APPROXIMATE SOLUTION

The main idea of the PKF is to approximate the conditional pdfs in Eq. (2) by some Gaussian mixture models (GMM), so that the solution to the data assimilation problem, in terms of some conditional pdfs, are also expressed in terms of some GMMs. For deduction, readers are referred to [5, 6]. Similar works can also be found in, for example, [8, 9, 10]. The similarity and difference between these existing works and ours were pointed out in [6], while the present study is different from our previous work [6] in that here we use both the stochastic EnKF [1] and its deterministic counterpart, the ensemble square root filter (EnSRF) [2, 3, 4], to construct the PKF, together with a customized pdf re-approximation scheme similar to that used in [6].

Due to the space limitation, we are not able to report the full details of the PKF in this work. Instead, we choose to give an outline of the procedures in the PKF.

1. Initialization: Given the initial background ensemble, one constructs a prior GMM which captures the mean and covariance of the initial background ensemble. Each individual Gaussian distribution in the GMM is associated with a background ensemble of the system state, which has the same (sample) background mean and covariance as those specified for the Gaussian distribution.
2. Filtering step: For each background ensemble associated with a Gaussian distribution, when an incoming observation is available, update the background ensemble to its analysis counterpart, based on a certain EnKF method, e.g., either the stochastic EnKF or the EnSRF. With the analysis ensembles, one is able to compute the (sample) analysis mean and covariance for each Gaussian distribution. Moreover, the weights associated with the Gaussian

distributions are also updated based on Bayes' rule. The mean and covariance of the GMM is then the weighted average of the means and covariances of individual Gaussian distributions in the GMM.

3. PDF re-approximation: Similar to the situation in the particle filter (PK), it is likely that the weights associated with the Gaussian distributions may collapse onto one single distribution. To prevent this, one may conduct "re-sampling" as if the Gaussian distributions were the particles in the PF. After "re-sampling", while the weights of the Gaussian distributions in the new GMM are set to be uniform, the new Gaussian distributions themselves may not be the same as those in the previous GMM. Hence the new analysis ensembles associated with the Gaussian distributions in the new GMM may also be different. Essentially, the new GMM is a new pdf different from the original GMM.
4. Propagation step: Propagate each new analysis ensemble forward to obtain the corresponding background ensemble at the next assimilation cycle, just as in a stochastic EnKF or an EnSRF. After that, repeat the procedures in the filtering step as aforementioned.

We argue that, in the case that there is only one Gaussian distribution in the GMM, the PKF reduces to the stochastic EnKF or the EnSRF. The stochastic EnKF can also be considered as an approximate solution of the data assimilation problem even when the pdf of the system state is non-Gaussian (e.g., there exist multiple Gaussian distributions in the GMM), which, however, is not true for the EnSRF.

NUMERICAL RESULTS

We choose the 40-dimensional system model due to Lorenz and Emanuel [11, 12] (LE98 model hereafter) as the testbed. The governing equations are given by

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + 8, \quad i = 1, \dots, 40. \quad (3)$$

For consistency, we define $x_{-1} = x_{39}$, $x_0 = x_{40}$, and $x_{41} = x_1$.

To generate synthetic observations, we adopt the following quadratic observation system

$$\mathbf{y}_k = 0.05(x_{k,1}^2, \dots, x_{k,39}^2) + \mathbf{v}_k, \quad (4)$$

where only the odd state variables $x_{k,i}$ ($i = 1, 3, \dots, 39$) of the system state $\mathbf{x}_k \equiv (x_{k,1}, \dots, x_{k,40})$ at time index k are observed, and the observation noise \mathbf{v}_k follows the 20-dimensional Gaussian distribution $N(\mathbf{v}_k : \mathbf{0}, \mathbf{I}_{20})$ with \mathbf{I}_{20} being the 20×20 identity matrix.

We use a fourth-order Runge-Kutta method to integrate (and discretize) the LE98 model from time 0 to 35, with a constant integration step of 0.05. To avoid the effect of transition states, we discard the trajectory from 0 to 25, while it is the period from 25.05 to 35 that is used for data assimilation. The observations are made for every four steps (corresponding to a time of 0.2).

To measure the performance of a filter, we calculate the time-averaged root mean squared error (rmse for short). In general, given a set of m_x -dimensional state vectors $\{\mathbf{x}_k : \mathbf{x}_k = (x_{k,1}, \dots, x_{k,m_x}), k = 0, \dots, k_{max}\}$, with k_{max} being the maximum time index ($k_{max} = 199$ in our experiments), suppose that one has an analysis $\hat{\mathbf{x}}_k^a = (\hat{x}_{k,1}^a, \dots, \hat{x}_{k,m_x}^a)$ of \mathbf{x}_k , then the rmse \hat{e} is defined as

$$\hat{e} = \frac{1}{k_{max} + 1} \sum_{k=0}^{k_{max}} e_k, \quad (5)$$

$$e_k \equiv \|\hat{\mathbf{x}}_k^a - \mathbf{x}_k\|_2 / \sqrt{m_x} = \sqrt{\frac{1}{m_x} \sum_{i=1}^{m_x} (\hat{x}_{k,i}^a - x_{k,i})^2},$$

where $\|\bullet\|_2$ denotes the Euclidean norm in \mathbb{R}^{m_x} such that $\|\mathbf{x}_k\|_2 = \sqrt{\sum_{j=1}^{m_x} x_{k,j}^2}$.

A possible problem in directly using \hat{e} as the performance measure is that \hat{e} itself may depend on some intrinsic parameters of the filters, for instance, the covariance inflation factor δ as to be discussed later, which might lead to inconsistent conclusions at different parameter values. To avoid this problem, we may sometimes adopt the following strategy: we relate a filter's best possible performance to the minimum rmse \hat{e}_{min} , which is the minimum value of \hat{e} that the filter can achieve within the chosen ranges of the filter's intrinsic parameters. In performance comparison, if

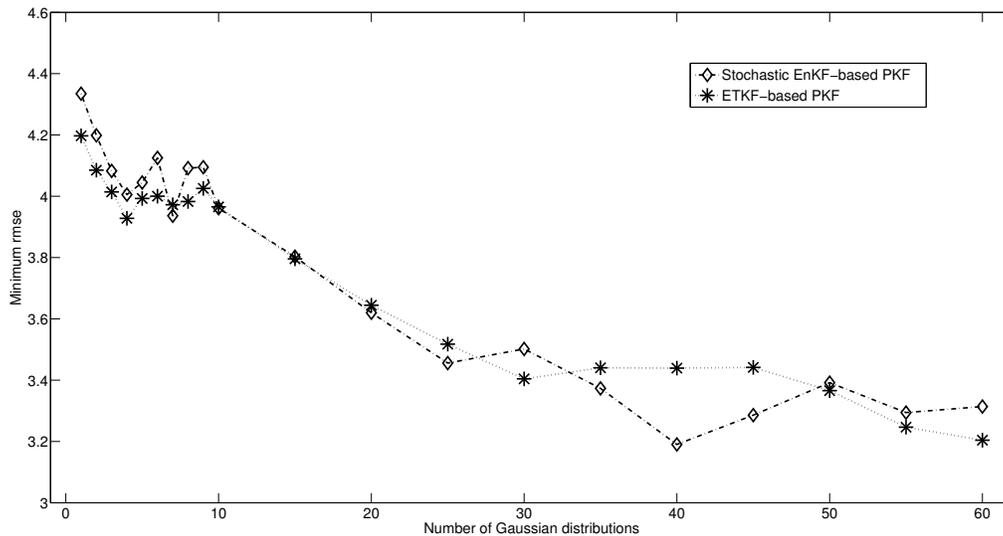


FIGURE 1. Minimum rms errors \hat{e}_{min} (over 20 experiments) of the stochastic EnKF- and ETKF-based PKFs (with a fixed ensemble size of 10 in each ensemble filter) as the function of the number of Gaussian pdfs in the GMM.

the minimum rmse \hat{e}_{min}^A of filter A is less than the minimum rmse \hat{e}_{min}^B of filter B , then we say that A performs better than B .

The experiment settings are as follows. We initialize the GMM with q Gaussian pdfs, where q takes the integer values from 1 to 10, with an even increment of 1 each time, and the integer values from 15 to 60, with an even increment of 5 each time. For convenience, we denote this choice by $q \in \{1 : 1 : 10, 15 : 5 : 60\}$, where the notation $v_{min} : v_{inc} : v_{max}$ represents a set of values which increases from v_{min} to v_{max} , with an even increment of v_{inc} each time. We choose to conduct pdf re-approximation, which introduces a real parameter, called the *fraction coefficient* c ($0 < c < 1$), to the PKF. To also examine its effect on the performance of the PKF, in our experiment we let $c \in \{0.05 : 0.1 : 0.95\}$. We use both the stochastic EnKF [1] and the ensemble transform Kalman filter (ETKF) [3], one of the EnSRFs, to construct the PKF. We let the ensemble size $n = 10$ in each type of the EnKF, which is relatively small compared to the system dimension 40. In this case, it is customary to conduct covariance inflation [8, 4] and localization [13] to make the PKF become more robust and perform better [14, 15]. In our experiments we let the covariance inflation factor $\delta = 0.02$. We follow the settings in Luo et al. [6, § 7.2.3] to conduct covariance localization and choose the length scale $l_c = 50$. To reduce statistical fluctuations, we repeat the experiments for 20 times, each time with a randomly drawn initial background ensemble, but the same true trajectory and the corresponding observations.

We examine the minimum rms errors \hat{e}_{min} of the stochastic EnKF- or ETKF-based PKFs within the tested values of c and q . In Fig. 1, we plot \hat{e}_{min} of the standard PKFs as functions of the number q of the Gaussian distributions, which shows that \hat{e}_{min} in either the stochastic EnKF- or ETKF-based PKF tends to decrease as the number q of Gaussian distributions increases. Note that $q = 1$ corresponds to the case that reduces to the single EnKF (the stochastic EnKF or the ETKF) itself. Therefore Fig. 1 implies that the PKF with $q > 1$ in general performs better than the corresponding ETKF, which confirms the benefit in adopting the PKF for data assimilation.

REFERENCES

1. G. Evensen, *J. Geophys. Res.* **99**(C5), 10143–10162 (1994).
2. J. L. Anderson, *Mon. Wea. Rev.* **129**, 2884–2903 (2001).
3. C. H. Bishop, B. J. Etherton, and S. J. Majumdar, *Mon. Wea. Rev.* **129**, 420–436 (2001).
4. J. S. Whitaker, and T. M. Hamill, *Mon. Wea. Rev.* **130**, 1913–1924 (2002).
5. I. Hoteit, D.-T. Pham, G. Triantafyllou, and G. Korres, *Mon. Wea. Rev.* **136**, 317–334 (2008).
6. X. Luo, I. M. Moroz, and I. Hoteit, *Physica D* **239**, 684–701 (2010).

7. M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, *IEEE Transactions on Signal Processing* **50**, 174–188 (2002).
8. J. L. Anderson, and S. L. Anderson, *Mon. Wea. Rev.* **127**, 2741–2758 (1999).
9. T. Bengtsson, C. Snyder, and D. Nychka, *J. Geophys. Res.* **108**, 8775 (2003).
10. K. W. Smith, *Tellus* **59A**, 749–757 (2007).
11. E. N. Lorenz, “Predictability—a problem partly solved,” in *Predictability*, edited by T. Palmer, ECMWF, Reading, UK, 1996.
12. E. N. Lorenz, and K. A. Emanuel, *J. Atmos. Sci.* **55**, 399–414 (1998).
13. T. M. Hamill, J. S. Whitaker, and C. Snyder, *Mon. Wea. Rev.* **129**, 2776–2790 (2001).
14. T. M. Hamill, J. S. Whitaker, J. L. Anderson, and C. Snyder, *Journal of the Atmospheric Sciences* **66**, 3498–3500 (2009).
15. P. J. Van Leeuwen, *Mon. Wea. Rev.* **137**, 4089–4114 (2009).