Jan Mandel; Loren Cobb; Jonathan D. Beezley
On the convergence of the ensemble Kalman filter

Applications of Mathematics, Vol. 56 (2011), No. 6, 533–541

Persistent URL: `http://dml.cz/dmlcz/141765`

ON THE CONVERGENCE OF THE ENSEMBLE KALMAN FILTER*

Jan Mandel, Loren Cobb, Jonathan D. Beezley, Denver

(Received September 29, 2009)

Abstract. Convergence of the ensemble Kalman filter in the limit for large ensembles to the Kalman filter is proved. In each step of the filter, convergence of the ensemble sample covariance follows from a weak law of large numbers for exchangeable random variables, the continuous mapping theorem gives convergence in probability of the ensemble members, and L^p bounds on the ensemble then give L^p convergence.

Keywords: data assimilation, ensemble, asymptotics, convergence, filtering, exchangeable random variables

MSC 2010: 62M20, 93E11

1. Introduction

Data assimilation uses statistical estimation to update the state of a running model based on new data. Data assimilation is of great importance and widely used in many disciplines including numerical weather prediction [10], ocean modeling [7], remote sensing [17], and image reconstruction [8]. In these applications, the dimension of the state is very high, often millions and more, because the state consists of the values of a simulation on a computational grid in a spatial domain. Consequently, the classical Kalman filter (KF), which requires maintaining the state covariance matrix, is no longer feasible.

One of the most successful recent data assimilation methods for high-dimensional problems is the ensemble Kalman filter (EnKF). EnKF is a Monte Carlo approximation of the KF, with the covariance in the KF replaced by the sample covariance computed from an ensemble of realizations. Because the EnKF does not need to

* This work was partially supported by the National Science Foundation under grants CNS-0719641 and AGS-0835579 and the National Institutes of Health under grant 1 RC1 LM01641-01.

maintain the state covariance matrix, it can be implemented efficiently for high-dimensional problems. Although the EnKF formulas rely on the assumption that the distribution of the state and the data likelihood are normal, the ensemble can robustly describe an arbitrary state probability distribution. Thus, in spite of errors such as smearing of the state distribution towards normality [13], the EnKF is often used for nonlinear systems.

One of the reasons for the popularity of the EnKF in applications is that the convergence of EnKF with the ensemble size tends to be quite fast and reasonably small ensembles (typically 25 to 100) are usually sufficient [7]. Convergence of the EnKF can be further accelerated by localization, such as covariance tapering [9], which improves the accuracy of the sample covariance. The EnKF converges rapidly in these applications because the state vectors are not arbitrary; rather, they are discretizations of smooth functions on a spatial domain, and so they are the states of an infinitely dimensional dynamical system. One explanation is that the state moves along a low-dimensional attractor. Indeed, in weather simulations, the EnKF performance can be further improved by a carefully chosen initial ensemble, which approximately covers the attractor well [10]. Another explanation is that a smooth random field can be well approximated by a linear combination of a small number of smooth functions with random coefficients, such as a truncated random Fourier series or Karhunen-Loève expansion. Indeed, if the state is not smooth enough, the convergence of the EnKF deteriorates [3] and large ensembles would be needed for acceptable accuracy.

A large body of literature on the EnKF and variants exists, but rigorous probabilistic analysis is lacking. It is commonly assumed that the ensemble is a sample (that is, i.i.d.) and that it is normally distributed. Although the resulting analyses played an important role in the development of EnKF, both assumptions are false. The ensemble covariance is computed from all ensemble members together, thus introducing dependence, and the EnKF formula is a nonlinear function of the ensemble, thus destroying the normality of the ensemble distribution.

For example, the analysis in [5] is based on the comparison of the covariance of the analysis ensemble and the covariance of the filtering distribution. The paper [9] notes that if the ensemble sample covariance is a consistent estimator, then Slutsky's theorem yields the convergence in probability of the gain matrix. The paper [12] studies the interplay of numerical and stochastic errors. All of these analyses assume that the ensemble covariance converges in some sense in the limit for large ensembles, but a rigorous justification has not yet become available.

This paper provides a rigorous proof that the EnKF converges to the KF in the limit for large ensembles and for normal state probability distributions and normal data likelihoods. The present analysis does not assume that the ensemble members are independent or normally distributed. The ensemble members are shown to be exchangeable random variables bounded in all L^p, $p \in [1, \infty)$, which provides properties that replace independence and normality. An argument using uniform integrability and the continuous mapping theorem is then possible.

The result is valid for the EnKF version of Burgers, van Leeuven, and Evensen [5] in the case of constant state space dimension, a linear model, normal data likelihood and initial state distributions, and ensemble size going to infinity. This EnKF version involves randomization of data. Efficient variants of EnKF without randomization exist [2], [15], but they are not the subject of this paper.

Probabilistic analysis of the performance of the EnKF on nonlinear systems, for non-normal state probability distributions, as well as the analysis of the speed of convergence of the EnKF to the KF and the dependence of the required ensemble size on the state dimension, are outside of the scope of this paper and left to future research. Some computational experiments and heuristic explanations can be found in [3].

After the original preprint of this paper was completed [14], some related work became available. The proof of EnKF convergence in [6] has a gap; it assumes that certain covariances derived from the ensemble exist, which is not guaranteed without an L^2 bound. The proof in [11] is related and also uses a priori L^p bounds, but it appears to be much longer and more complicated in order to obtain further analysis.

2. Preliminaries

The Euclidean norm of column vectors in $\mathbb{R}^m$, $m \geqslant 1$, and the induced matrix norm are denoted by $\|\cdot\|$, and $^\top$ is the transpose. The stochastic L^p norm of a random element X is $\|X\|_p = (E(\|X\|^p))^{1/p}$. The jth entry of a vector X is $[X]_j$ and the i, j entry of a matrix $Y \in \mathbb{R}^{m\times n}$ is $[Y]_{ij}$. Convergence in probability is denoted by $\xrightarrow{\mathrm{P}}$. We denote by

$$X_N = [X_{Ni}]_{i=1}^N = [X_{N1}, \ldots, X_{NN}],$$

with various superscripts and for various $m \geqslant 1$, an ensemble of N random elements in $\mathbb{R}^m$, called members. Thus, an ensemble is a random $m \times N$ matrix with the ensemble members as columns. Given two ensembles X_N and Y_N, the stacked ensemble $[X_N; Y_N]$ is defined as the block random matrix

$$[X_N; Y_N] = \begin{bmatrix} X_N \\ Y_N \end{bmatrix} = \left[\begin{bmatrix} X_{N1} \\ Y_{N1} \end{bmatrix}, \ldots, \begin{bmatrix} X_{NN} \\ Y_{NN} \end{bmatrix} \right] = [X_{Ni}; Y_{Ni}]_{i=1}^N.$$

If all the members of X_N are identically distributed, we write $E(X_{N1})$ and $\operatorname{Cov}(X_{N1})$ for their common mean vector and covariance matrix. The ensemble sample mean and ensemble sample covariance matrix are the random elements $\overline{X}_N = N^{-1}\sum_{i=1}^{N} X_{Ni}$ and $C(X_N) = \overline{X_N X_N^\top} - \overline{X}_N \overline{X}_N^\top$. *All convergence is for* $N \to \infty$.

We will work with ensembles such that the joint distribution of the ensemble X_N is invariant under permutation of the ensemble members. Such an ensemble is called *exchangeable*. That is, an ensemble X_N, $N \geqslant 2$, is exchangeable if and only if $\Pr(X_N \in B) = \Pr(X_N \Pi \in B)$ for every Borel set $B \subset \mathbb{R}^{m\times N}$ and every permutation matrix $\Pi \in \mathbb{R}^{N\times N}$. The covariance between any two members of an exchangeable ensemble is the same, $\operatorname{Cov}(X_{Ni}, X_{Nj}) = \operatorname{Cov}(X_{N1}, X_{N2})$, if $i \neq j$.

Lemma 1. *Suppose* X_N *and* D_N *are exchangeable, the random elements* X_N *and* D_N *are independent, and* $Y_{Ni} = F(X_N, X_{Ni}, D_{Ni})$, $i = 1, \ldots, N$, *where* F *is measurable and permutation invariant in the first argument, i.e.* $F(X_N\Pi, X_{Ni}, D_{Ni}) = F(X_N, X_{Ni}, D_{Ni})$ *for any permutation matrix* Π. *Then* Y_N *is exchangeable.*

P r o o f. Write $Y_N = \mathbf{F}(X_N, D_N)$, where

$$\mathbf{F}(X_N, D_N) = [F(X_N, X_{N1}, D_{N1}), F(X_N, X_{N2}, D_{N2}), \ldots, F(X_N, X_{NN}, D_{NN})].$$

Let Π be a permutation matrix. Then $Y_N\Pi = \mathbf{F}(X_N\Pi, D_N\Pi)$. Because X_N is exchangeable, the distributions of X_N and $X_N\Pi$ are identical. Similarly, the distributions of D_N and $D_N\Pi$ are identical. Since X_N and D_N are independent, the joint distributions of (X_N, D_N) and $(X_N\Pi, D_N\Pi)$ are identical. Thus, for any Borel set $B \subset \mathbb{R}^{n\times N}$,

$$\begin{aligned}\Pr(Y_N\Pi \in B) &= E(1_B(Y_N\Pi)) = E(1_B(\mathbf{F}(X_N\Pi, D_N\Pi))) \\ &= E(1_B(\mathbf{F}(X_N, D_N))) = \Pr(Y_N \in B),\end{aligned}$$

where 1_B stands for the characteristic function of B. Hence, Y_N is exchangeable. $\square$

We now prove a weak law of large numbers for nearly i.i.d. exchangeable ensembles.

Lemma 2. *If for all* N, X_N, U_N *are ensembles of random variables,* $[X_N; U_N]$ *is exchangeable,* $\operatorname{Cov}(U_{Ni}, U_{Nj}) = 0$ *for all* $i \neq j$, $U_{N1} \in L^2$ *is the same for all* N, *and* $X_{N1} \to U_{N1}$ *in* L^2, *then* $\overline{X}_N \xrightarrow{\mathrm{P}} E(U_{N1})$.

P r o o f. Since X_N is exchangeable, $\operatorname{Cov}(X_{Ni}, X_{Nj}) = \operatorname{Cov}(X_{N1}, X_{N2})$ for all $i, j = 1, \ldots, N$, $i \neq j$. Since $X_N - U_N$ is exchangeable, also $X_{N2} - U_{N2} \to 0$ in L^2. Then, using the identity $\operatorname{Cov}(X, Y) = E(XY) - E(X)E(Y)$ and the Cauchy inequality for the L^2 inner product $E(XY)$, we have

$$\begin{aligned}&|\operatorname{Cov}(X_{N1}, X_{N2}) - \operatorname{Cov}(U_{N1}, U_{N2})| \\ &\qquad \leqslant 2\|X_{N1}\|_2 \|X_{N2} - U_{N2}\|_2 + 2\|U_{N2}\|_2 \|X_{N1} - U_{N1}\|_2,\end{aligned}$$

so $\operatorname{Cov}(X_{N1}, X_{N2}) \to 0$. By the same argument, $\operatorname{Var}(X_{N1}) \to \operatorname{Var}(U_{N1}) < \infty$. Now $E(\overline{X}_N) = E(X_{N1}) \to E(U_{N1})$ from $X_{N1} - U_{N1} \to 0$ in L^2, and

$$\begin{aligned}\operatorname{Var}(\overline{X}_N) &= \frac{1}{N^2} \sum_{i=1}^{N} \operatorname{Var}(X_{Ni}) + \sum_{i,j=1, i\neq j}^{N} \operatorname{Cov}(X_{Ni}, X_{Nj}) \\ &= \frac{1}{N} \operatorname{Var}(X_{N1}) + \Big(1 - \frac{1}{N}\Big) \operatorname{Cov}(X_{N1}, X_{N2}) \to 0,\end{aligned}$$

and the conclusion follows from the Chebyshev inequality. $\square$

The convergence of the ensemble sample covariance follows.

Lemma 3. *If for all N, X_N, U_N are ensembles of random elements in $\mathbb{R}^n$, $[X_N; U_N]$ is exchangeable, U_N are i.i.d., $U_{N1} \in L^4$ is the same for all N, and $X_{N1} \to U_{N1}$ in L^4, then $\overline{X}_N \xrightarrow{\mathrm{P}} E(U_{N1})$ and $C(X_N) \xrightarrow{\mathrm{P}} \operatorname{Cov}(U_{N1})$.*

P r o o f. From Lemma 2, it follows that $[\overline{X}_N]_j \xrightarrow{\mathrm{P}} [E(U_{N1})]_j$ for each entry $j = 1, \ldots, n$, so $\overline{X}_N \xrightarrow{\mathrm{P}} E(U_{N1})$. Let $Y_{Ni} = X_{Ni} X_{Ni}^\top$, so that $C(X_N) = \overline{Y}_N - \overline{X}_N \overline{X}_N^\top$. Each entry of $[Y_{Ni}]_{jl} = [X_{Ni}]_j [X_{Ni}]_l$ satisfies the assumptions of Lemma 2, so $[Y_{Ni}]_{jl} \xrightarrow{\mathrm{P}} E([U_{N1} U_{N1}^\top]_{jl})$. Convergence of the entries $[\overline{X}_N \overline{X}_N^\top]_{jl} = [\overline{X}_N]_j [\overline{X}_N]_l$ to $E([U_{N1}]_{jl}) E([U_{N1}^\top]_{jl})$ follows from the already proved convergence of $\overline{X}_N$ and the continuous mapping theorem [16, p. 7]. Applying the continuous mapping theorem again, we get $C(X_N) \xrightarrow{\mathrm{P}} \operatorname{Cov}(U_{N1})$. $\square$

3. Formulation of the EnKF

Consider an initial state given as the random variable $U^{(0)}$. In step k, the state $U^{(k-1)}$ is advanced in time by applying the model $M^{(k)}$ to obtain $U^{(k),f} = M^{(k)}(U^{(k-1)})$, called the prior or the forecast, with probability density function (pdf) $p_{U^{(k),f}}$. The data in step k are given as measurements $d^{(k)}$ with a known error distribution, and expressed as the data likelihood $p(d^{(k)}|u)$. The new state $U^{(k)}$ conditional on the data, called the posterior or the analysis, then has the density $p_{U^{(k)}}$ given by the Bayes theorem,

$$p_{U^{(k)}}(u) \propto p(d^{(k)}|u) p_{U^{(k),f}}(u),$$

where $\propto$ means proportional. This is the discrete-time filtering problem. The distribution of $U^{(k)}$ is called the filtering distribution.

Assume that $U^{(0)} \sim N(u^{(0)}, Q^{(0)})$, the model is linear, $M^{(k)}\colon\ u \mapsto A^{(k)}u + b^{(k)}$, and the data likelihood is normal conditional on given state $u^{(k),f}$,

$$p(d^{(k)}|u^{(k),f}) \propto \mathrm{e}^{(-1/2)(H^{(k)}u^{(k),f}-d^{(k)})^\top {R^{(k)}}^{-1}(H^{(k)}u^{(k),f}-d^{(k)})},$$

where $H^{(k)}$ is the given observation matrix and $R^{(k)}$ is the given data error covariance. The data error is assumed to be independent of the model state. Then the filtering distribution is normal, $U^{(k)} \sim N(u^{(k)}, Q^{(k)})$, and it satisfies the KF recursions [1]

$$u^{(k),f} = E(U^{(k),f}) = A^{(k)}u^{(k)} + b^{(k)}, \quad Q^{(k),f} = \operatorname{Cov} U^{(k),f} = {A^{(k)}}^\top Q^{(k)} A^{(k)}, \tag{3.1}$$

$$u^{(k)} = u^{(k),f} + K^{(k)}(d^{(k)} - H^{(k)}u^{(k),f}), \quad Q^{(k)} = (I - K^{(k)}H^{(k)})Q^{(k),f}, \tag{3.2}$$

where the Kalman gain matrix $K^{(k)}$ is given by

$$K^{(k)} = Q^{(k),f}H^{(k)\top}(H^{(k)}Q^{(k),f}H^{(k)\top} + R^{(k)})^{-1}. \tag{3.3}$$

The EnKF is obtained by replacing the exact covariance $Q^{(k)}$ by the ensemble sample covariance and adding noise to the data in order to avoid a shrinking of the ensemble spread and to obtain the correct filtering covariance [5], cf. Lemma 4 below.

Let $U_i^{(0)} \sim N(u^{(0)}, Q^{(0)})$ and $D_i^{(k)} \sim N(d^{(k)}, R^{(k)})$ be independent for all $k, i \geqslant 1$. Given N, choose the initial ensemble and the perturbed data as the first N terms of the respective sequence, $U_{Ni}^{(0)} = U_i^{(0)}$, $i = 1, \ldots, N$, $D_{Ni}^{(k)} = D_i^{(k)}$, $i = 1, \ldots, N$, $k = 1, 2, \ldots$ The ensembles produced by EnKF are $X_N^{(0)} = U_N^{(0)}$ and

$$X_{Ni}^{(k),f} = M^{(k)}(X_{Ni}^{(k-1)}), \quad i = 1, \ldots, N, \tag{3.4}$$

$$X_N^{(k)} = X_N^{(k),f} + K_N^{(k)}(D_N^{(k)} - H^{(k)}X_N^{(k),f}), \tag{3.5}$$

where $K_N^{(k)}$ is the ensemble sample gain matrix,

$$K_N^{(k)} = Q_N^{(k),f}H^{(k)\top}(H^{(k)}Q_N^{(k),f}H^{(k)\top} + R^{(k)})^{-1}, \quad Q_N^{(k),f} = C(X_N^{(k),f}). \tag{3.6}$$

Our analysis of the EnKF is based on the observation that the ensembles $X_N^{(k)}$ are a perturbation of auxiliary ensembles $U_N^{(k)}$. The ensembles $U_N^{(k)}$ are obtained from the same initial ensemble by applying the KF formulas to each ensemble member separately and using the same corresponding member of perturbed data,

$$U_{Ni}^{(k),f} = M^{(k)}(U_{Ni}^{(k-1)}), \quad i = 1, \ldots, N, \tag{3.7}$$

$$U_N^{(k)} = U_N^{(k),f} + K^{(k)}(D_N^{(k)} - H^{(k)}U_N^{(k),f}). \tag{3.8}$$

The auxiliary ensembles $U_N^{(k)}$ are introduced for theoretical purposes only and they do not play any role in the EnKF algorithm. The next lemma shows that $U_N^{(k)}$ is a sample from the filtering distribution.

Lemma 4. *For all* $k = 1, 2, \ldots$, $U_N^{(k)}$ *is i.i.d. and* $U_{N1}^{(k)} \sim N(u^{(k)}, Q^{(k)})$.

P r o o f. The statement is true for $k=0$ by definition of $U_N^{(0)}$. Assume that it is true for $k-1$ in place of k. The ensemble $U_N^{(k)}$ is i.i.d. and normally distributed, because it is the image under a linear map of the normally distributed i.i.d. ensemble with members $[U_{Ni}^{(k-1)}, D_{Ni}^{(k)}]$, $i=1,\ldots,N$. Further, $D_N^{(k)}$ and $U_{Ni}^{(k),f}$ are independent, so from [5, Eq. (15) and (16)], $U_{N1}^{(k)}$ has the correct mean and covariance, which uniquely determines the normal distribution of $U_{N1}^{(k)}$. $\square$

4. Convergence analysis

Lemma 5. *There exist constants $c(k,p)$ for all k and all $p \in [1,\infty)$ such that $\|X_{Ni}^{(k)}\|_p \leqslant c(k,p)$ and $\|K_N^{(k)}\|_p \leqslant c(k,p)$ for all N.*

P r o o f. For $k=0$, each $X_{Ni}^{(k)}$ is normal. Assume $\|X_{Ni}^{(k-1)}\|_p \leqslant c(k-1,p)$ for all N. Then

$$\|X_{Ni}^{(k),f}\|_p = \|A^{(k)}X_{Ni}^{(k-1)} + b^{(k)}\|_p \leqslant \|A^{(k)}\|\|X_{Ni}^{(k-1)}\|_p + \|b^{(k)}\| \leqslant \mathrm{const}(k,p).$$

By Jensen's inequality, for any X_N,

$$\left\|\frac{1}{N}\sum_{i=1}^{N} X_{Ni}\right\|_p \leqslant \frac{1}{N}\sum_{i=1}^{N}\|X_{Ni}\|_p.$$

This gives $\|\overline{X}_N^{(k),f}\|_p \leqslant \mathrm{const}(k,p)$ and

$$\begin{aligned}\|Q_N^{(k),f}\|_p &\leqslant \frac{1}{N}\|X_{N1}^{(k),f}X_{N1}^{(k),f\top}\|_p + \frac{1}{N^2}\|X_{N1}^{(k),f}\|_p^2 \\ &\leqslant \frac{1}{N}\|X_{N1}^{(k),f}\|_{2p}^2 + \frac{1}{N^2}\|X_{N1}^{(k),f}\|_p^2 \leqslant \mathrm{const}(k,p),\end{aligned}$$

since from the Cauchy inequality

$$\begin{aligned}\|WZ\|_p &\leqslant E(\|W\|^p\|Z\|^p)^{1/p} \\ &\leqslant E(\|W\|^{2p})^{1/(2p)}E(\|Z\|^{2p})^{1/(2p)} = \|W\|_{2p}\|Z\|_{2p},\end{aligned} \tag{4.1}$$

for any compatible random matrices W and Z. Since $H^{(k)}Q_N^{(k),f}H^{(k)\top}$ is symmetric positive semidefinite and $R^{(k)}$ is symmetric positive definite, it holds that

$$\|(H^{(k)}Q_N^{(k),f}H^{(k)\top} + R^{(k)})^{-1}\| \leqslant \|(R^{(k)})^{-1}\| \leqslant \mathrm{const}(k),$$

which, together with the bound on $\|Q_N^{(k),f}\|_p$, gives

$$\|K_N^{(k)}\|_p \leqslant \|Q_N^{(k)}\|_p\,\mathrm{const}(k) \leqslant \mathrm{const}(k,p).$$

Finally, we obtain the desired bound

$$\begin{aligned}\|X_{Ni}^{(k)}\|_p &\leqslant \|X_{Ni}^{(k),f}\|_p + \|K_N^{(k)} D_{Ni}^{(k)}\|_p + \|K_N^{(k)} H^{(k)} X_{Ni}^{(k),f}\|_p \\ &\leqslant \operatorname{const}(k,p)(\|X_{Ni}^{(k),f}\|_p + \|K_N^{(k)}\|_p + \|K_N^{(k)}\|_{2p}\|X_{Ni}^{(k),f}\|_{2p}) \leqslant c(k,p),\end{aligned}$$

using again (4.1). □

Theorem 1. *For all k, $[X_N; U_N]$ is exchangeable and $X_{Ni}^{(k)} \to U_{Ni}^{(k)}$ in L^p for all $p \in [1,\infty)$, where U_N is i.i.d. with the filtering distribution.*

Proof. The ensembles $U_N^{(k)}$ are obtained by linear mapping of the i.i.d. initial ensemble $U_N^{(0)}$, so they are i.i.d. For $k = 1$, we have $X_N^{(0)} = U_N^{(0)}$, $[X_N^{(0)}; U_N^{(0)}]$ is exchangeable, and $X_{Ni} = U_{Ni}$. Suppose the statement holds for $k-1$ in place of k. The ensemble members are given by a recursion of the form

$$[X_{Ni}^{(k)}; U_{Ni}^{(k)}] = F^{(k)}(C(X_N^{(k-1)}), [X_{Ni}^{(k-1)}; U_{Ni}^{(k-1)}], D_{Ni}^{(k)}).$$

The ensemble sample covariance matrix $C(X_N^{(k-1)})$ is invariant to a permutation of ensemble members, so $[X_N^{(k)}; U_N^{(k)}]$ is exchangeable by Lemma 1. Since $X_N^{(k),f}$ and $U_N^{(k),f}$ satisfy the assumptions of Lemma 3, it follows that $C(X_N^{(k),f}) \xrightarrow{\mathrm{P}} \operatorname{Cov} U_{N1}^{(k),f}$ and $K_N^{(k)} \xrightarrow{\mathrm{P}} K^{(k)}$. Thus, comparing (3.5) and (3.8), we have that $X_{Ni}^{(k)} \xrightarrow{\mathrm{P}} U_{Ni}^{(k)}$, by the continuous mapping theorem. Let $p \in [1,\infty)$. Since the sequence $\{X_{Ni}^{(k)}\}_{N=1}^{\infty}$ is bounded in L^p by Lemma 5 and $X_{Ni}^{(k)} \xrightarrow{\mathrm{P}} U_{Ni}^{(k)}$, it follows that $X_{Ni}^{(k)} \to U_{Ni}^{(k)}$ in L^q for all $1 \leqslant q < p$ by uniform integrability [4, p. 338]. □

Using Lemma 3 and uniform integrability again, it follows that the ensemble mean and covariance converge to the filtering mean and covariance.

Corollary 1. *$\overline{X}_N^{(k)} \to u^{(k)}$ and $C(X_N^{(k)}) \to Q^{(k)}$ in L^p for all $p \in [1,\infty)$, where $u^{(k)}$ and $Q^{(k)}$ are the mean and the covariance of the filtering distribution.*

Acknowledgement. We would like to thank Professor Richard Bradley for an example that led us away from mixing coefficients, and Dr. Thomas Bengtsson for bringing [9] to our attention.

References

[1] *B. D. O. Anderson, J. B. Moore*: Optimal Filtering. Prentice-Hall, Englewood Cliffs, 1979.

[2] *J. L. Anderson*: An ensemble adjustment Kalman filter for data assimilation. Monthly Weather Review *129* (1999), 2884–2903.

[3] *J. D. Beezley*: High-dimensional data assimilation and morphing ensemble Kalman filters with applications in wildfire modeling. PhD. Thesis. Department of Mathematical and Statistical Sciences, University of Colorado Denver, 2009, http://math.cudenver.edu/~jbeezley/jbeezley_thesis.pdf.

[4] *P. Billingsley*: Probability and Measure, 3rd edition. John Wiley & Sons Inc., New York, 1995.

[5] *G. Burgers, P. J. van Leeuwen, G. Evensen*: Analysis scheme in the ensemble Kalman filter. Monthly Weather Review *126* (1998), 1719–1724.

[6] *M. D. Butala, R. A. Frazin, Y. Chen, F. Kamalabadi*: Tomographic imaging of dynamic objects with the ensemble Kalman filter. IEEE Trans. Image Proces. *18* (2009), 1573–1587.

[7] *G. Evensen*: Data Assimilation: The Ensemble Kalman Filter. Springer, Berlin, 2007.

[8] *R. A. Frazin, M. D. Butala, A. Kemball, F. Kamalabadi*: Time-dependent reconstruction of nonstationary objects with tomographic or interferometric measurements. Astrophys. J. *635* (2005), L197–L200.

[9] *R. Furrer, T. Bengtsson*: Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. J. Multivariate Anal. *98* (2007), 227–255.

[10] *E. Kalnay*: Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press, Cambridge, 2003.

[11] *F. Le Gland, V. Monbet, V.-D. Tran*: Large Sample Asymptotics for the Ensemble Kalman Filter. Research Report RR-7014, INRIA, August 2009. http://hal.inria.fr/inria-00409060/en/.

[12] *J. Li, D. Xiu*: On numerical properties of the ensemble Kalman filter for data assimilation. Comput. Methods Appl. Mech. Eng. *197* (2008), 3574–3583.

[13] *J. Mandel, J. D. Beezley*: An ensemble Kalman-particle predictor-corrector filter for non-Gaussian data assimilation. ICCS 2009. Lecture Notes in Computer Science, Vol. 5545 (G. Allen, J. Nabrzyski, E. Seidel, G. D. van Albada, J. Dongarra, P. M. A. Sloot, eds.). Springer, 2009, pp. 470–478.

[14] *J. Mandel, L. Cobb, J. D. Beezley*: On the convergence of the ensemble Kalman filter. University of Colorado Denver CCM Report 278, January 2009. http://www.arXiv.org/abs/0901.2951.

[15] *M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill, J. S. Whitaker*: Ensemble square root filters. Monthly Weather Review *131* (2003), 1485–1490.

[16] *A. W. Van der Vaart*: Asymptotic Statistics. Cambridge University Press, Cambridge, 2000.

[17] *A. Vodacek, Y. Li, A. J. Garrett*: Remote sensing data assimilation in environmental models. In: AIPR '08: Proceedings of the 2008 37th IEEE Applied Imagery Pattern Recognition Workshop. IEEE Computer Society, Washington, 2008, pp. 1–5.

Authors' address: *J. Mandel* (corresponding author), *L. Cobb*, *J. D. Beezley*, Center for Computational Mathematics and Department of Mathematical and Statistical Sciences, University of Colorado Denver, Denver, CO 80217-3364, U.S.A., e-mail: Jan.Mandel@ucdenver.edu, Loren.Cobb@ucdenver.edu, Jonathan.Beezley@ucdenver.edu.