

# Convergence of the Ensemble Kalman Filter

Jan Mandel

Center for Computational Mathematics  
Department of Mathematical and Statistical Sciences  
University of Colorado Denver

Based on joint work with Loren Cobb and Jonathan Beezley  
<http://math.ucdenver.edu/~jmandel/papers/enkf-theory.pdf>  
<http://arxiv.org/abs/0901.2951>

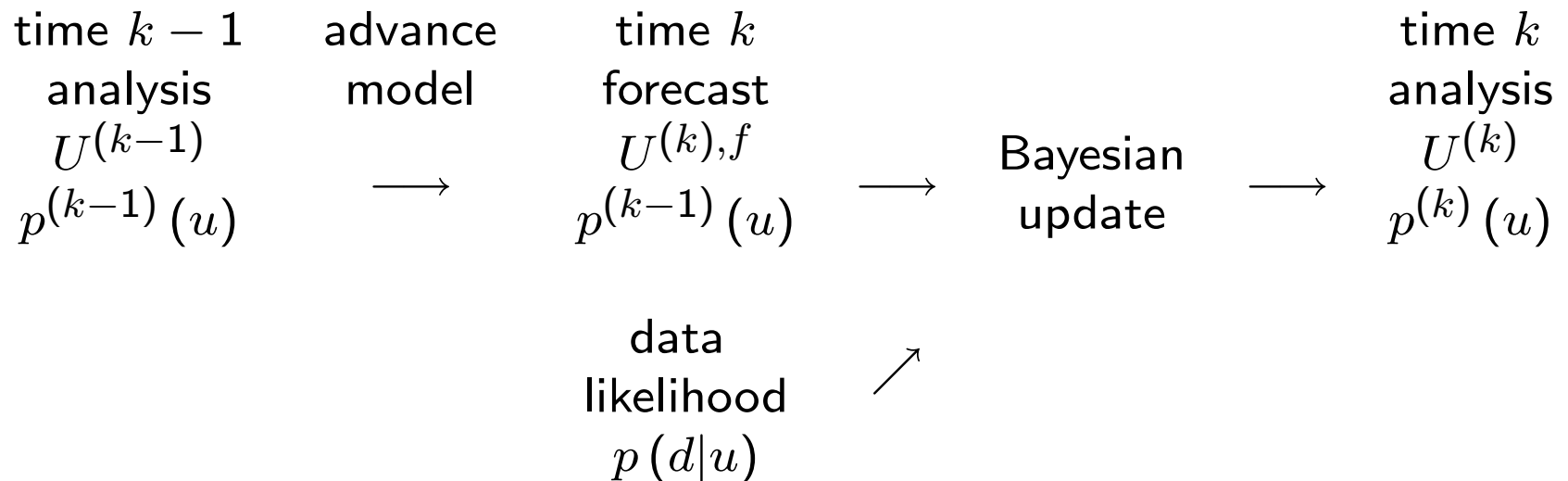
Supported by NSF grants CNS-071941 and ATM-0835579,  
and NIH grant 1 RC1 LM010641-01.

Department of Statistics, Colorado State University, October 5, 2009

# Data assimilation

- Goal: Inject data into a running model
- Approach: *discrete time filtering = sequential statistical estimation*
- Model *state is a random variable  $U$*  with probability density  $p(u)$
- Data + measurement error enter as *data likelihood  $p(d|u)$*
- Probability densities updated by *Bayes theorem*:  $p(u) \propto p(d|u) p_f(u)$

## Analysis cycle



## Kalman filter

- Model state  $U$  is represented by its mean and covariance.
- Assume that the forecast is Gaussian,  $U^f \sim N(m^f, Q^f)$ , and
- the data likelihood is Gaussian,

$$p(d|u) \propto \exp[-(Hu - d)^T R^{-1} (Hu - d) / 2].$$

- Then the analysis is also Gaussian,  $U \sim N(m, Q)$ , given by

$$m = m^f + K(d - Hm^f), \quad Q = (I - KH)Q^f$$

where  $K = Q^f H^T (HQ^f H^T + R)^{-1}$  is the gain matrix

- Hard to advance the covariance matrix when the model is nonlinear
- Needs to maintain the covariance matrix - unsuitable for large problems
- Ensemble Kalman filter (EnKF): represent  $U$  by an ensemble and estimate  $\text{Cov } U$  by the sample covariance.

# Ensemble Kalman filter (EnKF)

- represent model state by an ensemble, covariance  $\rightarrow$  sample covariance
- apply Kalman formulas for the mean to the ensemble members
- perturb data to get the correct analysis ensemble covariance

## Algorithm

1. Generate initial i.i.d. ensemble  $X_1, \dots, X_N$
2. In each analysis cycle

Advance the model in time: forecast  $X_i^f = \mathcal{M}(X_i)$

Compute the sample covariance  $Q_N^f$  of the forecast ensemble

Get randomized data:  $D_i \sim d + E_i$ , i.i.d.  $E_i \sim N(0, R)$

Bayesian update: analysis  $X_i = X_k^f + K_N(D_i - HX_i^f)$

where  $K_N = Q_N^f H^\top (H Q_N^f H^\top + R)^{-1}$

## EnKF properties

- The model is needed only as a black box.
- The EnKF was derived for Gaussian distributions.
- It is often used for nonlinear models and non-Gaussian distributions.
- The EnKF was designed so that the ensemble is a sample from the correct analysis distribution if
  - the forecast ensemble is i.i.d. and Gaussian
  - the covariance matrix is exact
- But
  - the sample covariance matrix is a random variable
  - the ensemble is not independent: sample covariance ties it together
  - the ensemble is not Gaussian: the update step is nonlinear

# Convergence analysis overview

- Goal: show that asymptotically, for large ensemble
  - the sample covariance converges to the exact covariance
  - the ensemble is asymptotically i.i.d.
  - the ensemble is asymptotically Gaussian
  - the ensemble members converge to the correct distribution
- Roadmap:
  - the ensemble is exchangeable: plays the role of independent
  - the ensemble is a-priori bounded in  $L^p$ : plays the role of Gaussian
  - use the weak law of large numbers for the sample covariance
  - the ensemble converges to a Gaussian i.i.d. ensemble obtained with the exact covariance matrix instead of the sample covariance
  - tools: weak convergence and  $L^p$  convergence of random variables

## Tools: Weak convergence

Weak convergence of probability measures:  $\mu_n \Rightarrow \mu$

if  $\int f d\mu_n \rightarrow \int f d\mu \forall$  continuous bounded  $f$

Weak convergence of random elements:  $X_n \Rightarrow X$

if  $\text{law}(X_n) \Rightarrow \text{law}(X)$

Continuous mapping theorem: if  $X_n \Rightarrow X$  and  $g$  is continuous,

then  $g(X_n) \Rightarrow g(X)$

Joint weak convergence: if  $X_n \Rightarrow X$  and  $Y_n \Rightarrow C$  (constant),

then  $(X_n, Y_n) \Rightarrow (X, C)$

Slutsky's theorem: if  $X_n \Rightarrow X$  and  $Y_n \Rightarrow C$  are random variables,

then  $X_n + Y_n \Rightarrow X + C$  and  $X_n Y_n \Rightarrow X C$

## Tools: $L^p$ convergence

Stochastic  $L^p$  norm of a random element  $X$  :  $\|X\|_p = (E(|X|^p))^{1/p}$

$L^p$  convergence implies weak convergence: if  $X_n \rightarrow X$  in  $L^p$ ,  
then  $X_n \Rightarrow X$

Uniform integrability: if  $X_n \Rightarrow X$  and  $\{X_n\}$  bounded in  $L^p$ ,  
then  $X_n \rightarrow X$  in  $L^q \forall q < p$

## Convergence to Kalman filter

- Generate the initial ensemble  $X_N^{(0)} = \{X_{N1}^{(0)}, \dots, X_{NN}^{(0)}\}$  i.i.d. and Gaussian.
- Run EnKF, get ensembles  $\{X_{N1}^{(k)}, \dots, X_{NN}^{(k)}\}$ .
- For theoretical purposes, define  $U_N^{(0)} = X_N^{(0)}$  and  $U_N^{(k)}$  by EnKF except with the exact covariance of the filtering distribution instead of the sample covariance.
- Then  $U_N^{(k)}$  is a sample (i.i.d.) from the Gaussian filtering distribution (as would be delivered by the Kalman filter).
- **We will show that  $X_{Ni}^{(k)} \rightarrow U_{Ni}^{(k)}$  in  $L^p$  for all  $p < \infty$  as  $N \rightarrow \infty$ , for all analysis cycles  $k$ .**

# A-priori properties of EnKF ensembles

## Exchangeability

**Definition:** set of r.v. is exchangeable if their joint distribution is permutation invariant.

**Lemma:** *All EnKF ensembles are exchangeable.*

*Proof:* The initial ensemble in EnKF is i.i.d., thus exchangeable, and the analysis step is permutation invariant.

# A-priori properties of EnKF ensembles

## $L^p$ bounds

**Lemma.** *All EnKF ensembles are bounded in  $L^p$  with constant independent of  $N$ , for all  $p < \infty$ .*

*Proof ingredients:*

- Recall that

$$X_N^{(k)} = X_N^{(k),f} + K_N^{(k)} (D_N^{(k)} - H^{(k)} X_N^{(k),f})$$
$$K_N^{(k)} = Q_N^{(k),f} H^{(k)T} (H^{(k)} Q_N^{(k),f} H^{(k)T} + R^{(k)})^{-1}$$

- Bound  $\| Q_N^{(k),f} \|_p$  from the  $L^{2p}$  bound on  $X_N^{(k),f}$
- Bound the matrix norm  $|(H^{(k)} Q_N^{(k),f} H^{(k)T} + R^{(k)})^{-1}| \leq \text{const}$  from  $R^{(k)} > 0$  (positive definite).

## Convergence of sample mean

**Lemma.** If  $\left\{ \begin{bmatrix} X_{N1} \\ U_{N1} \end{bmatrix}, \dots, \begin{bmatrix} X_{NN} \\ U_{NN} \end{bmatrix} \right\}$  are exchangeable,  $\{U_{N1}, \dots, U_{NN}\}$  are i.i.d.,  $U_{N1} \in L^2$  is the same for all  $N$ , and  $X_{N1} \rightarrow U_{N1}$  in  $L^2$ , then  $\bar{X}_N = (X_{N1} + \dots + X_{NN})/N \Rightarrow E(U_{N1})$ .

Note that  $\{U_{N1}, \dots, U_{NN}\}$  and  $\{X_{N1}, \dots, X_{NN}\}$  are not independent.

*Proof ingredients:*

Exchangeability gives  $\|X_{Ni} - U_{Ni}\|_2 \rightarrow 0$  and  $\text{Cov}(X_{Ni}, X_{Nj}) \rightarrow 0$  independent of  $i, j$ . Standard Monte-Carlo argument:

$$\begin{aligned} \|\bar{X}_N - E(X_{N1})\|^2 &= \text{Var}(\bar{X}_N) \\ &= \frac{1}{N^2} \sum_{i=1}^N \text{Var}(X_{Ni}) + \sum_{i,j=1, j \neq i}^N \text{Cov}(X_{Ni}, X_{Nj}) \\ &= \frac{1}{N} \text{Var}(X_{N1}) + \left(1 - \frac{1}{N}\right) \text{Cov}(X_{N1}, X_{N2}) \rightarrow 0 \end{aligned}$$

## Convergence of sample covariance

**Lemma.** *If  $\left\{ \begin{bmatrix} X_{N1} \\ U_{N1} \end{bmatrix}, \dots, \begin{bmatrix} X_{NN} \\ U_{NN} \end{bmatrix} \right\}$  are exchangeable,  $\{U_{N1}, \dots, U_{NN}\}$  are i.i.d.,  $U_{N1} \in L^4$  is the same for all  $N$ , and  $X_{N1} \rightarrow U_{N1}$  in  $L^4$ , then the sample covariance of  $(X_{N1}, \dots, X_{NN}) \Rightarrow \text{Cov}(U_{N1})$ ,  $N \rightarrow \infty$ .*

*Proof ingredients:*

Entries of the sample covariance are  $[\bar{X}_N \bar{X}_N^T]_{j\ell} - \overline{[X_{Ni}]_j [X_{Ni}]_\ell}$ .

$L^4$  bound on  $X_{Ni}$  gives  $L^2$  bound on the products  $[X_{Ni}]_j [X_{Ni}]_\ell$ , thus the sample mean of the products converges.

Use convergence of the sample mean  $\bar{X}_N$  and Slutsky's theorem.

## Recall EnKF ensemble and reference KF ensemble

Initial ensemble  $X_N^{(0)}$  is Gaussian i.i.d., then generate  $X_N^{(k)}$  from EnKF. Define  $U_N^{(0)} = X_N^{(0)}$  then  $U_N^{(k)}$  from EnKF, except with the exact KF covariance  $Q^{(k),f}$  instead of ensemble covariance  $Q_N^{(k),f}$ .

The model is linear:

$$X_{Ni}^{(k),f} = A^{(k)} X_{Ni}^{(k-1)} + b^{(k)}, \quad U_{Ni}^{(k),f} = A^{(k)} U_{Ni}^{(k-1)} + b^{(k)}$$

EnKF update:

$$\begin{aligned} X_{Ni}^{(k)} &= X_{Ni}^{(k),f} + K_N^{(k)} (D_{Ni}^{(k)} - H^{(k)} X_{Ni}^{(k),f}) \\ K_N^{(k)} &= Q_N^{(k),f} H^{(k)\top} (H^{(k)} Q_N^{(k),f} H^{(k)\top} + R^{(k)})^{-1} \end{aligned}$$

$$\begin{aligned} U_{Ni}^{(k)} &= U_{Ni}^{(k),f} + K^{(k)} (D_{Ni}^{(k)} - H^{(k)} U_{Ni}^{(k),f}) \\ K^{(k)} &= Q^{(k),f} H^{(k)\top} (H^{(k)} Q^{(k),f} H^{(k)\top} + R^{(k)})^{-1} \end{aligned}$$

Because  $Q^{(k),f}$  is exact,  $U_N^{(k)}$  is i.i.d. with the KF distribution.

## Proof of EnKF convergence to KF

**Theorem.**  $X_{Ni}^{(k)} \rightarrow U_{Ni}^{(k)}$  in  $L^p$  for all  $p < \infty$  as  $N \rightarrow \infty$ , for all  $k$ .

*Proof outline.* True for  $k = 0$ . Assume the statement is true for  $k - 1$ .

- Linear model gives  $X_{N1}^{(k),f} \rightarrow U_{N1}^{(k),f}$  in  $L^p$  as  $N \rightarrow \infty$ , for all  $p$ .
- Since  $X_N$  and  $U_N$  are jointly exchangeable, the sample covariance matrices converge weakly:  $Q_N^{(k),f} \Rightarrow Q^{(k),f}$ .
- By continuous mapping and Slutsky's theorem, the gain matrices also converge weakly:  $K_N^{(k)} \Rightarrow K_N$ .
- From Slutsky's theorem, then the analysis ensemble members converge weakly:  $X_{N1}^{(k)} \Rightarrow U_{N1}^{(k)}$ .
- From the a-priori  $L^p$  bound on  $X_N^{(k)}$  and uniform integrability, the EnKF ensemble members converge to the KF:  $X_{Ni}^{(k)} \rightarrow U_{Ni}^{(k)}$  in all  $L^q$ ,  $q < p$ .

## Conclusion

- We have proved convergence in large sample limit for arbitrary but fixed space dimension and number of analysis cycles, for EnKF with randomized data, and Gaussian state. For non-Gaussian state, EnKF is not justified as a Bayesian procedure anyway.

### **Open problems - future work and in progress**

- Convergence independent of large state dimension and convergence to KF on Hilbert space in progress.
- Long-time convergence, matrix Riccati equation.
- Frequent updates: convergence to a continuous time filter.
- Non-Gaussian case: But EnKF smears towards Gaussian. Theory for particle filter (PF) independent of dimension, and EnKF - particle filter hybrid.

## Convergence in high and infinite dimension (in progress)

- The real interest is in EnKF for high-dimensional problems. Our approach: Best considered as approximating infinite dimension.
- Solutions of PDE models lie naturally in spaces of smooth functions. Then the state distributions of discretized problems should approach a probability distribution on a space of functions (=random field).
- Kalman filter can be extended to Gaussian measures on Hilbert space. Proof using cylinder sets, apply matrix formulas on the finite dimensional base. Existing literature seems to deal with continuous time only.
- The EnKF asymptotics then mostly carries over to Hilbert space.
- The convergence of EnKF in finite dimension then should be bounded independently of dimension.
- Computational experiments confirm that EnKF converges uniformly for high-dimensional distributions that approximate a Gaussian measure on Hilbert space. (J. Beezley, Ph.D. thesis, 2009).
- Actually, the same is seen for particle filters... in spite of the fact that particle filters for problems without structure fail in high dimension.