Ensemble Methods in Geophysical Sciences
Météo France, Toulouse, November 12–16, 2012

Large sample asymptotics for the ensemble Kalman filter (EnKF)

François Le Gland
INRIA Rennes + IRMAR

Valérie Monbet
Université de Rennes 1 + IRMAR

supported by ANR project PREVASSEMBLE
coordinated by Olivier Talagrand

outline

- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- toy example
- central limit theorem for EnKF

non–linear state–space model of a special form

$$X_k \;=\; f_k(X_{k-1}) + W_k \qquad \text{with} \qquad W_k \sim \mathcal{N}(0, Q_k)$$

$$Y_k \;=\; H_k\, X_k + V_k \qquad \text{with} \qquad V_k \sim \mathcal{N}(0, R_k)$$

with non–necessarily Gaussian initial condition $X_0 \sim \eta_0$

observation noise covariance matrix R_k assumed invertible

in such a model, conditional probability distribution (aka Bayesian filter)

of hidden state X_k given past observations $Y_{0:k} = (Y_0, \cdots, Y_k)$

is not Gaussian

ensemble Kalman filter provides a Bayes–free Monte Carlo approach to numerically evaluate the Bayesian filter, as an alternative to particle filters

▶ initialization : initial ensemble $(X_0^{1,f}, \cdots, X_0^{N,f})$ is *simulated* as i.i.d. random vectors with probability distribution η_0, i.e. with same statistics as initial condition X_0

▶ prediction (forecast) step : given analysis ensemble $(X_{k-1}^{1,a}, \cdots, X_{k-1}^{N,a})$ each member is propagated independently according to (set of decoupled equations)

$$X_k^{i,f} = f_k(X_{k-1}^{i,a}) + W_k^i \qquad \text{with} \qquad W_k^i \sim \mathcal{N}(0, Q_k)$$

here, i.i.d. random vectors $(W_k^1, \cdots, W_k^N)$ are *simulated* with same statistics as additive Gaussian noise W_k in original state equation : in particular $(W_k^1, \cdots, W_k^N)$ are independent of forecast ensemble $(X_{k-1}^{1,a}, \cdots, X_{k-1}^{N,a})$

define empirical covariance matrix

$$P_k^{N,f} = \frac{1}{N} \sum_{i=1}^N (X_k^{i,f} - m_k^{N,f})\,(X_k^{i,f} - m_k^{N,f})^* \qquad \text{with} \qquad m_k^{N,f} = \frac{1}{N} \sum_{i=1}^N X_k^{i,f}$$

of forecast ensemble $(X_k^{1,f}, \cdots, X_k^{N,f})$

▶ correction (analysis) step : given forecast ensemble $(X_k^{1,f}, \cdots, X_k^{N,f})$ each member is updated independently according to (set of equations with mean–field interactions)

$$X_k^{i,a} = X_k^{i,f} + K_k(P_k^{N,f})\, (Y_k - H_k\, X_k^{i,f} - V_k^i) \qquad \text{with} \qquad V_k^i \sim \mathcal{N}(0, R_k)$$

in terms of Kalman gain mapping defined by

$$P \longmapsto K_k(P) = P\, H_k^* \, (H_k\, P\, H_k^* + R_k)^{-1}$$

for any $m \times m$ covariance matrix P,
and in terms of empirical covariance matrix $P_k^{N,f}$ of forecast ensemble

here, i.i.d. random vectors $(V_k^1, \cdots, V_k^N)$ are *simulated* with same statistics as additive Gaussian noise V_k in original observation equation : in particular $(V_k^1, \cdots, V_k^N)$ are independent of forecast ensemble $(X_k^{1,f}, \cdots, X_k^{N,f})$

mean–field interaction : given that

$$X_k^{i,a} = X_k^{i,f} + K_k(P_k^{N,f})\,(Y_k - H_k\, X_k^{i,f} - V_k^i)$$

each analysis member depends on whole forecast ensemble $(X_k^{1,f}, \cdots, X_k^{N,f})$ through empirical probability distribution

$$\mu_k^{N,f} = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_k^{i,f}}$$

actually only through empirical covariance matrix $P_k^{N,f}$ of forecast ensemble

$\longrightarrow$ results in *dependent* analysis ensemble $(X_k^{1,a}, \cdots, X_k^{N,a})$

question : does ensemble empirical probability distribution converge to Bayesian filter, defined as

$$\mu_k^-(dx) = \mathbb{P}[X_k \in dx \mid Y_{0:k-1}] \qquad \text{and} \qquad \mu_k(dx) = \mathbb{P}[X_k \in dx \mid Y_{0:k}]$$

i.e. does

$$\mu_k^{N,f} = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_k^{i,f}} \longrightarrow \mu_k^- \qquad \text{and} \qquad \mu_k^{N,a} = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_k^{i,a}} \longrightarrow \mu_k$$

hold in some sense, as $N \uparrow \infty$?

answer in general is negative

however, in the linear Gaussian case (i.e. if $f_k(x)$ depends linearly on x), then answer is positive and in particular ensemble empirical mean vector does converge to Kalman predictor / filter, i.e.

$$\widehat{X}_k^{N,f} = \frac{1}{N} \sum_{i=1}^{N} X_k^{i,f} \longrightarrow \widehat{X}_k^- \qquad \text{and} \qquad \widehat{X}_k^{N,a} = \frac{1}{N} \sum_{i=1}^{N} X_k^{i,a} \longrightarrow \widehat{X}_k$$

as $N \uparrow \infty$

decoupling approach : to study asymptotic behaviour of empirical probability distributions

$$\mu_k^{N,f} = \frac{1}{N} \sum_{i=1}^N \delta_{X_k^{i,f}} \qquad \text{and} \qquad \mu_k^{N,a} = \frac{1}{N} \sum_{i=1}^N \delta_{X_k^{i,a}}$$

of forecast ensemble and analysis ensemble, respectively, approximating i.i.d. random vectors are introduced as follows

initially $\bar{X}_0^{i,f} = X_0^{i,f}$, i.e. initial set of i.i.d. random vectors coincides exactly with initial forecast ensemble

these vectors are propagated independently according to (set of fully decoupled equations)

$$\bar{X}_k^{i,f} = f_k(\bar{X}_{k-1}^{i,a}) + W_k^i \qquad \text{with} \qquad W_k^i \sim \mathcal{N}(0, Q_k)$$

and

$$\bar{X}_k^{i,a} = \bar{X}_k^{i,f} + K_k(\bar{P}_k^f)\,(Y_k - H_k\,\bar{X}_k^{i,f} - V_k^i) \qquad \text{with} \qquad V_k^i \sim \mathcal{N}(0, R_k)$$

where $\bar{P}_k^f$ denotes (true) covariance matrix of i.i.d. random vectors $\bar{X}_k^{i,f}$

heuristics : these i.i.d. random vectors are close (contiguous) to members in ensemble Kalman filter, since they

- start from same initial values exactly
- use same i.i.d. random vectors $(W_k^1, \cdots, W_k^N)$ and $(V_k^1, \cdots, V_k^N)$ exactly, already *simulated* and used in ensemble Kalman filter

essentially a theoretical (not practical) concept

- large sample asymptotics is simple to analyze, because of independance
- true covariance matrix $\bar{P}_k^f$ is unknown, hence these i.i.d. random vectors are not computable in practice

in contrast, members in ensemble Kalman filter are computable but dependent, because they all contribute to / use empirical covariance matrix $P_k^{N,f}$ which results in mean–field interaction

outline

- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- toy example
- central limit theorem for EnKF

intuition : limiting probability distributions $\bar{\mu}_k^f$ and $\bar{\mu}_k^a$ are probability distributions of i.i.d. random vectors $\bar{X}_k^{i,f}$ and $\bar{X}_k^{i,a}$ respectively, and are completely characterized by integrals of arbitrary test functions

- initialization : recall that $\bar{X}_0^{i,f} = X_0^{i,f}$ and $X_0^{i,f} \sim \eta_0$, hence $\bar{\mu}_0^f = \eta_0$
- forecast (expression of $\bar{\mu}_k^f$ in terms of $\bar{\mu}_{k-1}^a$) : recall that

$$\bar{X}_k^{i,f} = f_k(\bar{X}_{k-1}^{i,a}) + W_k^i \qquad \text{with} \qquad W_k^i \sim \mathcal{N}(0, Q_k)$$

and since $\bar{X}_{k-1}^{i,a}$ has probability distribution $\bar{\mu}_{k-1}^a$ (by definition), then

$$\int_{\mathbb{R}^m} \phi(x')\, \bar{\mu}_k^f(dx') = \mathbb{E}[\phi(\bar{X}_k^{i,f})] = \mathbb{E}[\phi(f_k(\bar{X}_{k-1}^{i,a}) + W_k^i))]$$

$$= \int_{\mathbb{R}^m} \underbrace{\int_{\mathbb{R}^m} \phi(f_k(x) + w)\, p_k^W(dw)}_{T_k\,\phi(x)}\, \bar{\mu}_{k-1}^a(dx)$$

where $p_k^W(dw)$ is Gaussian probability distribution with zero mean vector and covariance matrix Q_k, i.e. probability distribution of random vector W_k^i

▶ analysis (expression of $\bar{\mu}_k^a$ in terms of $\bar{\mu}_k^f$) : recall that

$$\bar{X}_k^{i,a} = \bar{X}_k^{i,f} + K_k(\bar{P}_k^f)\,(Y_k - H_k\,\bar{X}_k^{i,f} - V_k^i) \qquad \text{with} \qquad V_k^i \sim \mathcal{N}(0, R_k)$$

sufficient conditions on drift function f_k can be given, under which $\bar{\mu}_k^f$ has finite second order moments, in which case covariance matrix $\bar{P}_k^f$ is finite

and since $\bar{X}_k^{i,f}$ has probability distribution $\bar{\mu}_k^f$ (by definition), then

$$\int_{\mathbb{R}^m} \phi(x')\,\bar{\mu}_k^a(dx') = \mathbb{E}[\phi(\bar{X}_k^{i,a})] = \mathbb{E}[\phi(\bar{X}_k^{i,f} + K_k(\bar{P}_k^f)\,(Y_k - H_k\,\bar{X}_k^{i,f} - V_k^i))]$$

$$= \int_{\mathbb{R}^m} \underbrace{\int_{\mathbb{R}^d} \phi(x + K_k(\bar{P}_k^f)\,(Y_k - H_k\,x - v))\,q_k^V(v)\,dv}_{T_k^{\mathrm{KF}}(\bar{\mu}_k^f)\,\phi(x)}\,\bar{\mu}_k^f(dx)$$

where $q_k^V(v)$ is Gaussian density with zero mean vector and invertible covariance matrix R_k, i.e. probability density of random vector V_k^i

on the other hand, Bayesian filter, defined as

$$\mu_k^-(dx) = \mathbb{P}[X_k \in dx \mid Y_{0:k-1}] \qquad \text{and} \qquad \mu_k(dx) = \mathbb{P}[X_k \in dx \mid Y_{0:k}]$$

satisfies recurrent relation

$$\int_{\mathbb{R}^m} \phi(x')\, \mu_k^-(dx') = \int_{\mathbb{R}^m} \underbrace{\int_{\mathbb{R}^m} \phi(f_k(x) + w)\, p_k^W(dw)}_{T_k\, \phi(x)}\, \mu_{k-1}(dx)$$

and (Bayes rule)

$$\int_{\mathbb{R}^m} \phi(x')\, \mu_k(dx') = \frac{\displaystyle\int_{\mathbb{R}^m} \phi(x')\, q_k^V(Y_k - H_k\, x')\, \mu_k^-(dx')}{\displaystyle\int_{\mathbb{R}^m} q_k^V(Y_k - H_k\, x')\, \mu_k^-(dx')}$$

with initial condition $\mu_0^- = \eta_0$

clearly, limiting probability distributions of forecast / analysis ensemble do not coincide with Bayesian predictor / filter, i.e. $\bar{\mu}_k^f \neq \mu_k^-$ and $\bar{\mu}_k^a \neq \mu_k$, except in the linear Gaussian case

outline

- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- toy example
- central limit theorem for EnKF

indeed, intuition is correct : ensemble empirical probability distributions

$$\mu_k^{N,\bullet} = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_k^{i,\bullet}}$$

do converge (in some sense) as $N \uparrow \infty$ to the probability distribution $\bar{\mu}_k^{\bullet}$ of i.i.d. random vectors $\bar{X}_k^{i,\bullet}$ (hence, not to the Bayesian filter)

reference

> F. Le Gland, V. Monbet and Vu–Duc Tran *Large sample asymptotics for the ensemble Kalman filter*, chapter 22 in The Oxford Handbook of Nonlinear Filtering, 2011

Theorem (law of large numbers) under mild assumptions on drift function f_k and on test function ϕ

$$\frac{1}{N} \sum_{i=1}^{N} \phi(X_k^{i,\bullet}) \longrightarrow \int_{\mathbb{R}^m} \phi(x)\, \bar{\mu}_k^{\bullet}(dx)$$

in probability as $N \uparrow \infty$

Theorem ($\mathbb{L}^p$–convergence and rate of convergence) under mild assumptions on drift function f_k and on test function ϕ, and provided initial condition X_0 has finite moments of any order p

$$\sup_{N \geq 1} \sqrt{N}\, (\, \mathbb{E}| \frac{1}{N} \sum_{i=1}^{N} \phi(X_k^{i,\bullet}) - \int_{\mathbb{R}^m} \phi(x)\, \bar{\mu}_k^{\bullet}(dx)\, |^p\,)^{1/p} < \infty$$

for any order p

to summarize : ensemble Kalman filter

- gain matrix depends on empirical covariance matrix
- ensemble empirical probability distribution converges to the wrong limit (different from Bayesian filter), except for linear Gaussian model
- rate of convergence $1/\sqrt{N}$

vs. (any brand of) particle filter

- weighted empirical probability distribution of particle system converges to the correct limit (Bayesian filter)
- rate of convergence $1/\sqrt{N}$, with central limit theorem

question : is there any advantage to use ensemble Kalman filter ?

idea : prove central limit theorem (and compare asymptotic error variances) outline

- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- toy example
- central limit theorem for EnKF

toy example

linear Gaussian system

the target distribution, i.e. the Bayesian filter, is known explicitly as a Gaussian distribution, with mean and covariance provided by the Kalman filter

hidden state

$$X_k = a\, X_{k-1} + \sqrt{1-a^2}\, W_k \qquad \text{with} \qquad W_k \sim \mathcal{N}(0, \sigma^2)$$

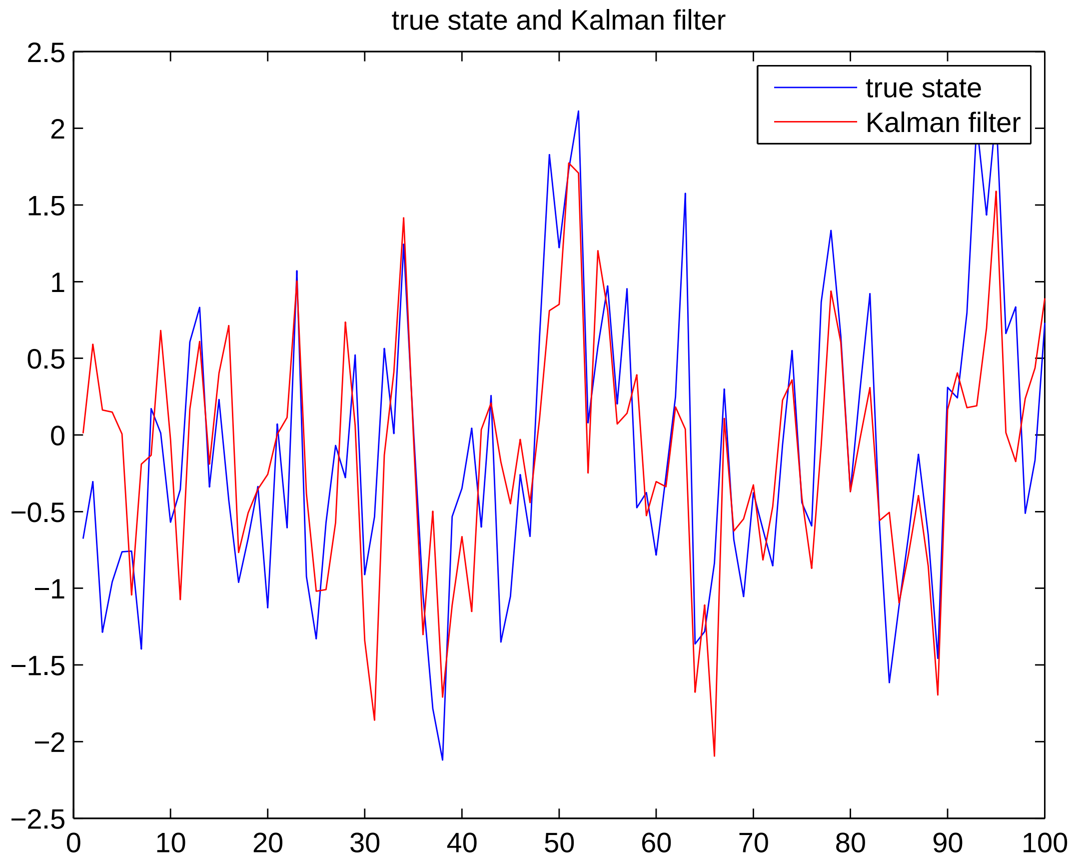
initial condition $X_0 \sim \mathcal{N}(0, \sigma^2)$ so that stationarity holds

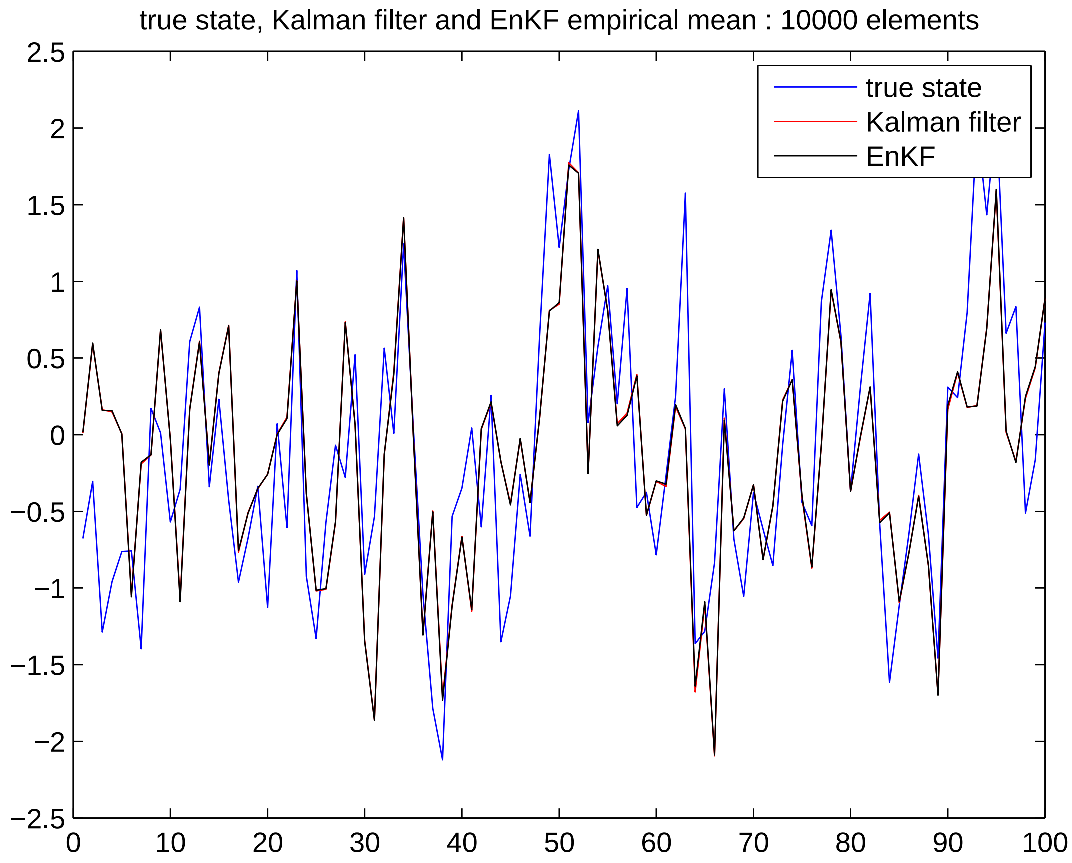
observations

$$Y_k = X_k + V_k \qquad \text{with} \qquad V_k \sim \mathcal{N}(0, s^2)$$

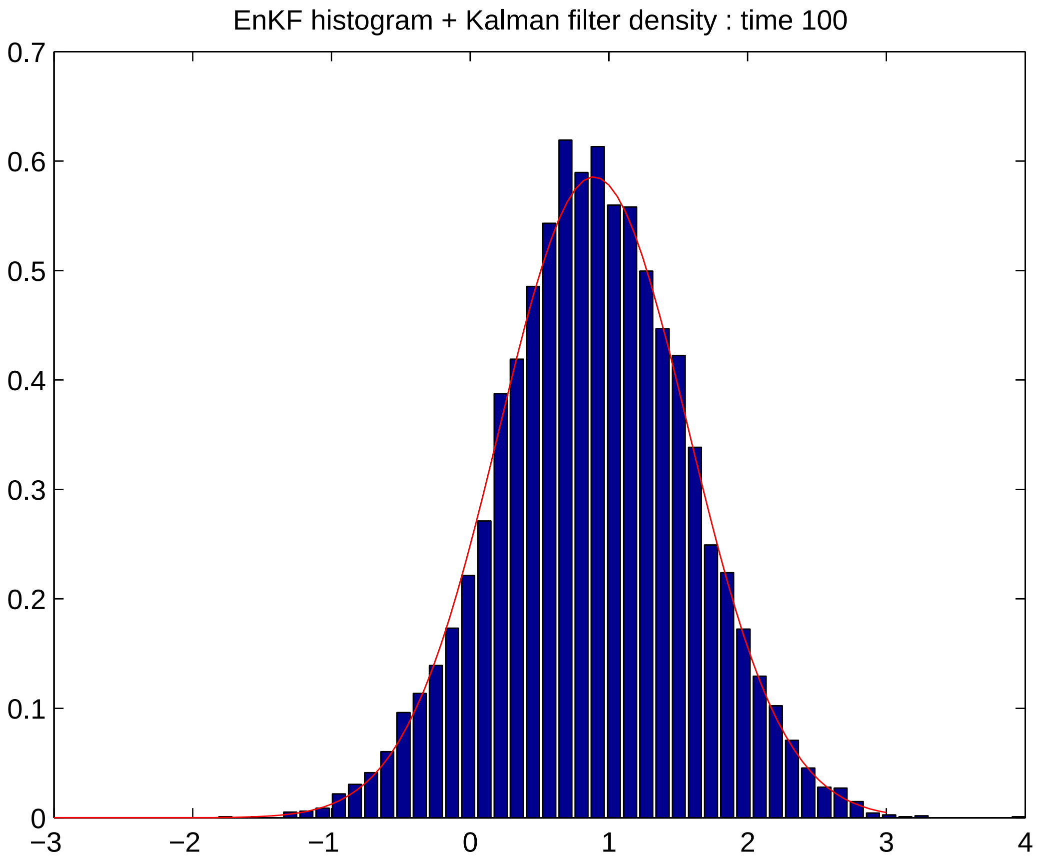
numerical values

a	σ	s
0.5	1	1


true state and Kalman filter
true state
Kalman filter
2.5
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2
−2.5
0
10
20
30
40
50
60
70
80
90
100



EnKF empirical mean vector $\widehat{X}_k^{N,a}$ with $N = 10000$ members



EnKF histogram with $N = 10000$ members

conclusion : not only does the EnKF empirical mean vector

$$\widehat{X}_k^{N,a} = \frac{1}{N} \sum_{i=1}^N X_k^{i,a}$$

converge to the Kalman filter $\widehat{X}_k$, but more generally the EnKF empirical probability distribution

$$\mu_k^{N,a} = \frac{1}{N} \sum_{i=1}^N \delta_{X_k^{i,a}}$$

converges to the Gaussian distribution with moments given by the Kalman filter

next different question : how fast does the empirical mean vector converge to the Kalman filter, e.g. is the normalized difference

$$\sqrt{N}\, (\widehat{X}_k^{N,a} - \widehat{X}_k) = \frac{1}{\sqrt{N}} \sum_{i=1}^N (X_k^{i,a} - \widehat{X}_k)$$

asymptotically normally distributed and how to compute the asymptotic variance ?

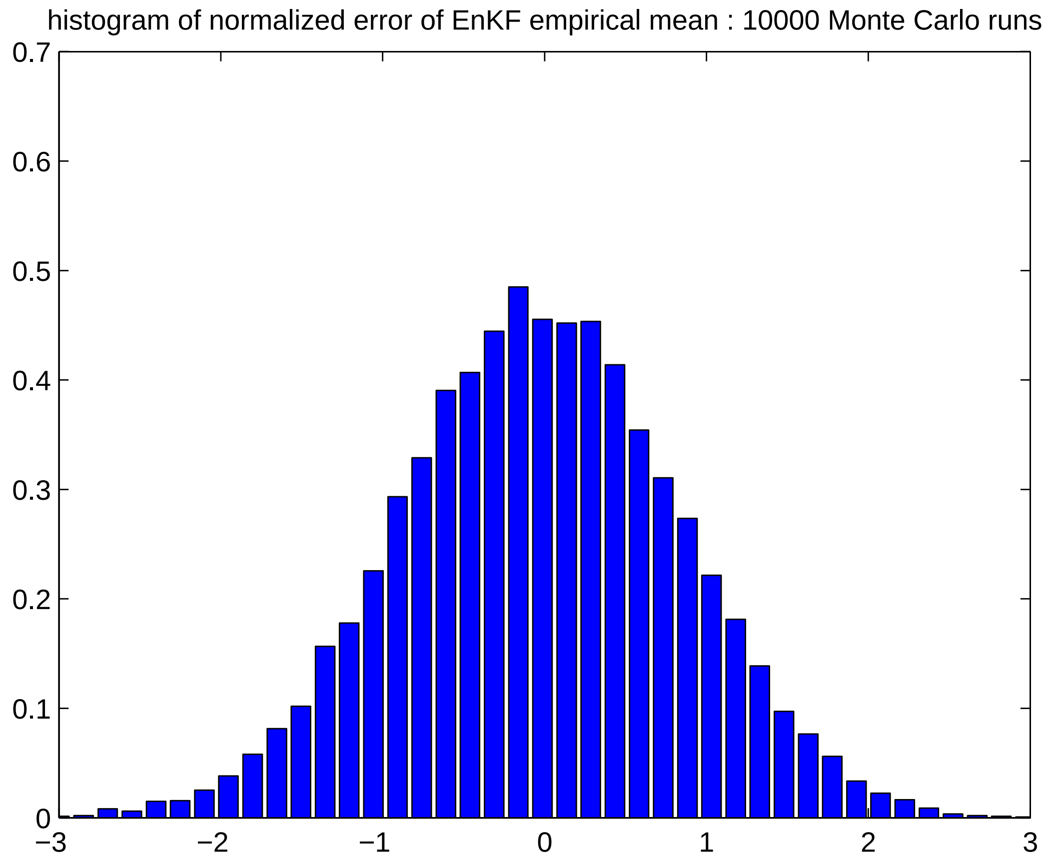
toy example (continued)

numerical simulations : for EnKF / bootstrap particle filter / particle filter with optimal importance distribution

- M Monte Carlo runs
- each Monte Carlo run evaluates one ensemble / particle average, based on N members / particles and compares this average with the (known) limit
- histogram of the M normalized differences is shown

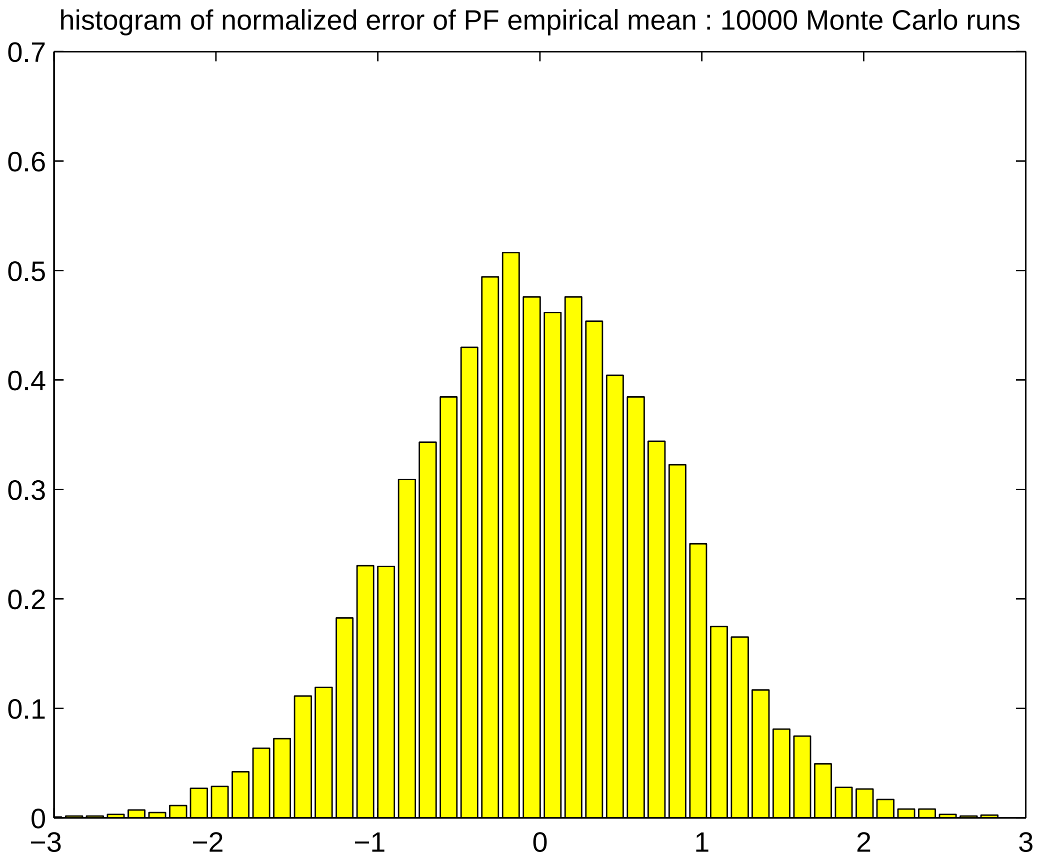
same toy example : stationary linear Gaussian system
with same numerical values

a	σ	s
0.5	1	1



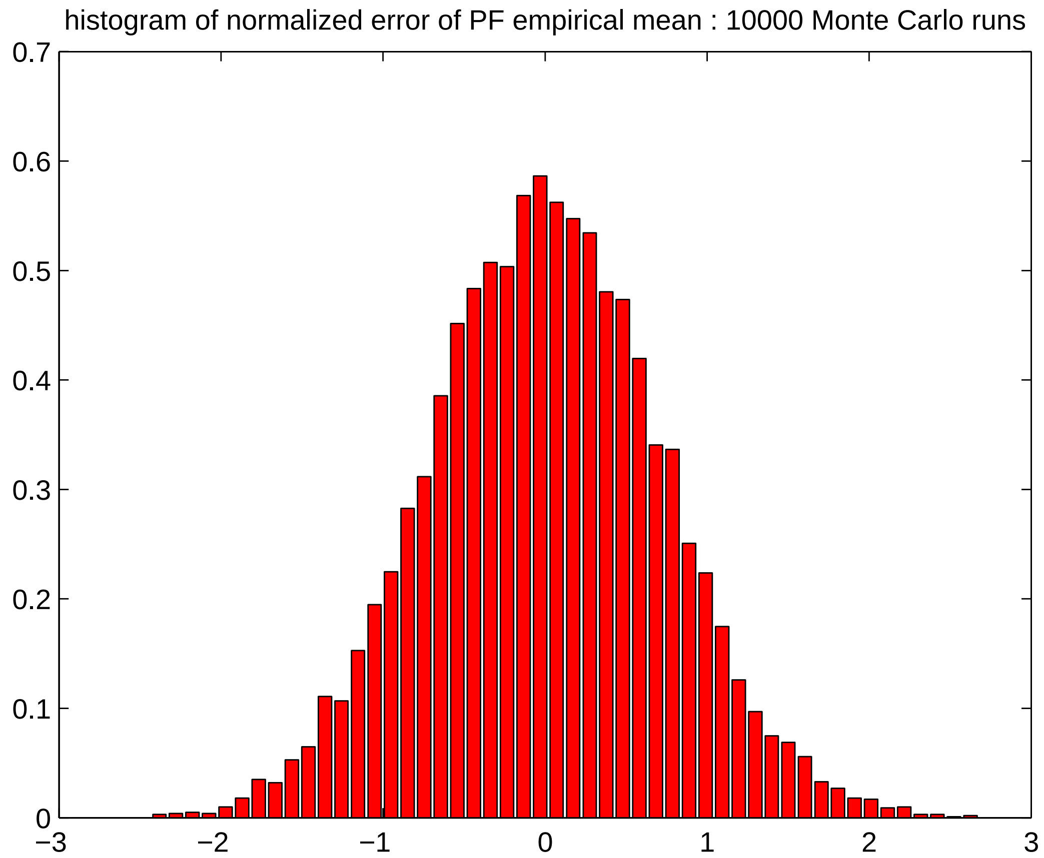
histogram of EnKF normalized differences $\sqrt{N}\,(\widehat{X}_k^{N,a} - \widehat{X}_k)$ for $N = 1000$ members and $M = 10000$ Monte Carlo runs

empirical standard deviation 0.880



histogram of (bootstrap) particle filter normalized differences $\sqrt{N}\;(\widehat{X}_k^N - \widehat{X}_k)$ for $N = 1000$ particles and $M = 10000$ Monte Carlo runs

empirical standard deviation 0.822



histogram of (optimal) particle filter normalized differences $\sqrt{N}\ (\widehat{X}_k^N - \widehat{X}_k)$ for $N = 1000$ particles and $M = 10000$ Monte Carlo runs

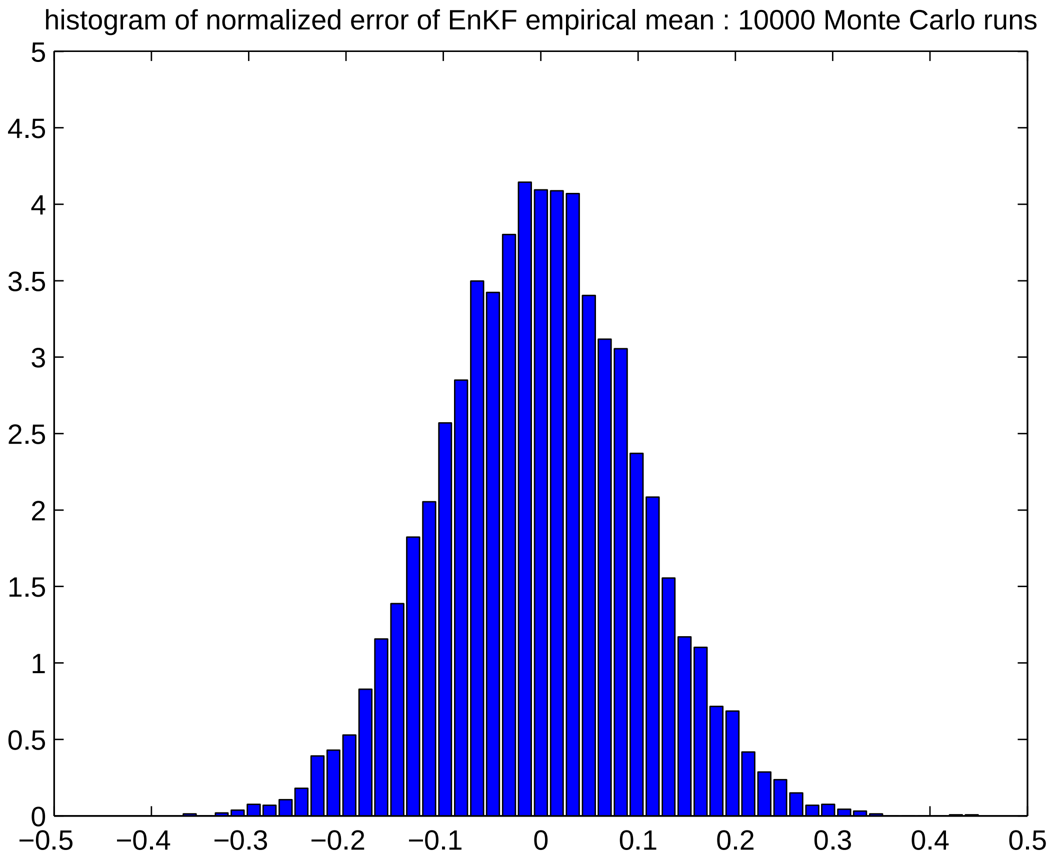
empirical standard deviation 0.713

first findings (based on these first simulations) : in terms of speed of convergence (a smaller asymptotic variance means a faster convergence)

PF with optimal importance distribution $\gg$ bootstrap PF $\gg$ EnKF

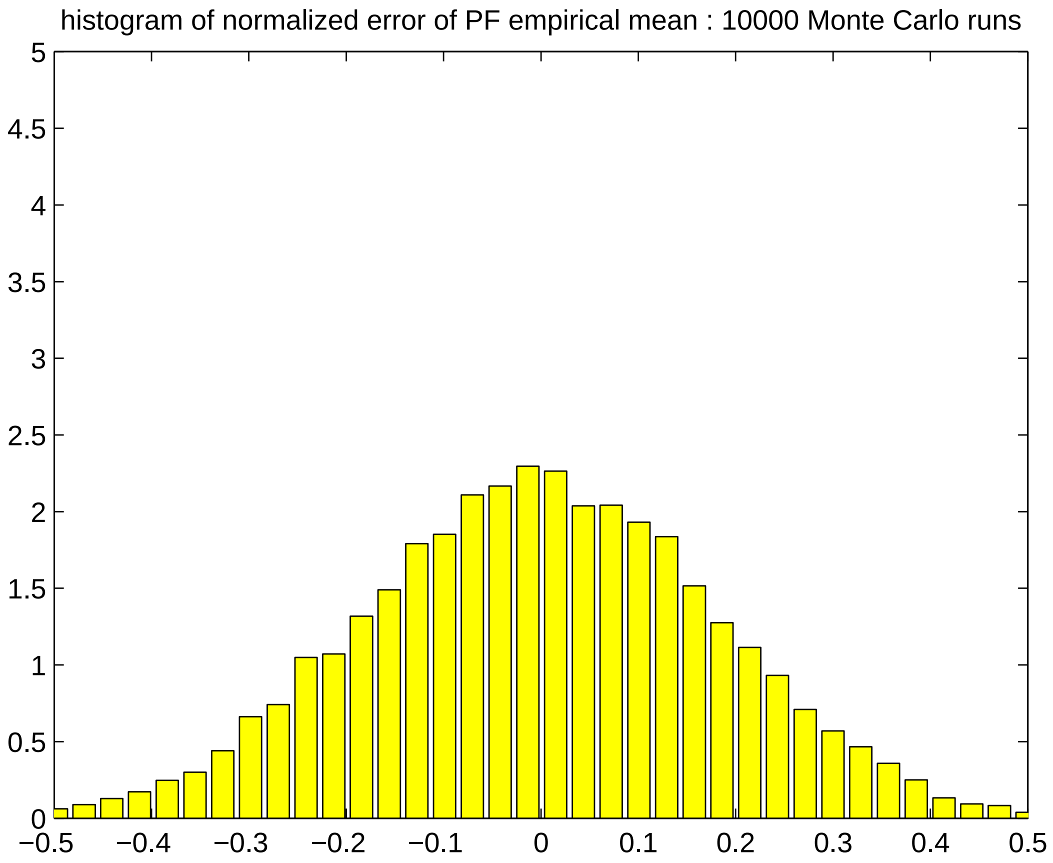
however, consider same toy example : stationary linear Gaussian system with different numerical values (smaller observation noise)

a	σ	s
0.5	1	0.01



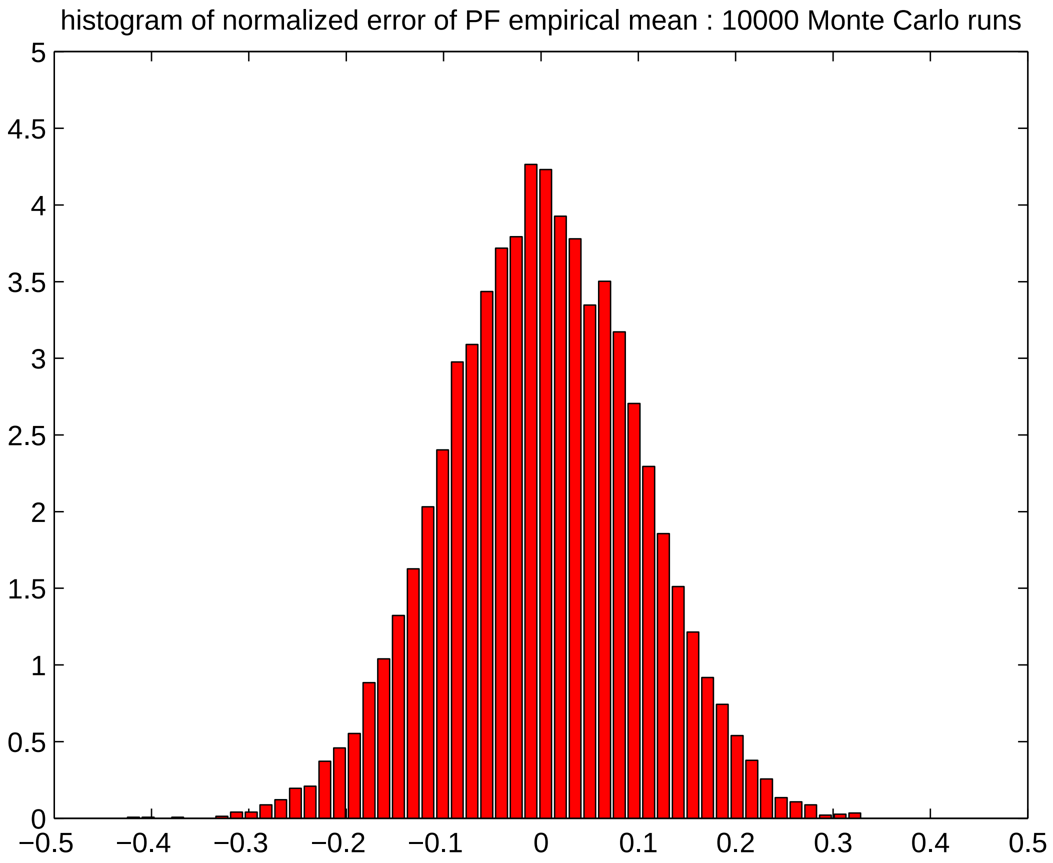
histogram of EnKF normalized differences $\sqrt{N}\,(\widehat{X}_k^{N,a} - \widehat{X}_k)$ for $N = 1000$ members and $M = 10000$ Monte Carlo runs

empirical standard deviation 0.100



histogram of (bootstrap) particle filter normalized differences $\sqrt{N}\ (\widehat{X}_k^N - \widehat{X}_k)$ for $N = 1000$ particles and $M = 10000$ Monte Carlo runs

empirical standard deviation 0.182



histogram of (optimal) particle filter normalized differences $\sqrt{N}\,(\widehat{X}_k^N - \widehat{X}_k)$ for $N = 1000$ particles and $M = 10000$ Monte Carlo runs

empirical standard deviation 0.099

somehow different findings (based on these different simulations) : in terms of speed of convergence (a smaller asymptotic variance means a faster convergence)

PF with optimal importance distribution $\approx$ EnKF $\gg$ bootstrap PF

conclusion : results have been obtained in the large sample asymptotics

- EnKF is (asymptotically) biased, does not converge to the optimal Bayesian filter, except in the linear Gaussian case
- in particular, empirical mean of EnKF ensemble does not converge to MMSE (conditional mean) of hidden state given past observations
- normalized approximation error (difference of empirical mean of EnKF ensemble and its limit) is asymptotically Gaussian, with (more or less computable) expression for the asymptotic variance

are these results relevant / can they provide any help or insight in the more practical case of a finite (small) ensemble size ?

outline

- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- toy example
- central limit theorem for EnKF

Theorem (central limit theorem) under mild assumptions on drift function f_k and on test function ϕ

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\phi(X_k^{i,\bullet}) - \int_{\mathbb{R}^m} \phi(x) \, \bar{\mu}_k^\bullet(dx)\,] \Longrightarrow \mathcal{N}(0, \bar{v}_k^\bullet(\phi))$$

in distribution as $N \uparrow \infty$, with (more or less explicit) expression for asymptotic variance $\bar{v}_k^\bullet(\phi)$

beyond the qualitative statement

- recurrence relations for the asymptotic variance ?
- practical computations ?

because of the recursive nature of the ensemble Kalman filter, it seems natural to prove the CLT by induction, and to rely on a strategy already used in Künsch (Annals of Statistics, 2005)

Lemma if

- conditionally w.r.t. $\mathcal{F}_N$, the r.v. Z'_N converges in distribution to a Gaussian r.v. with zero mean and variance V', in the sense that for any fixed u

$$\mathbb{E}[\exp\{j\, u\, Z'_N\} \mid \mathcal{F}_N] \longrightarrow \exp\{-\tfrac{1}{2}\, u^2\, V'\}$$

  in probability, and in $\mathbb{L}^1$ by the Lebesgue dominated convergence thorem

- the r.v. Z''_N is measurable w.r.t. $\mathcal{F}_N$, and converges in distribution to a Gaussian r.v. with zero mean and variance V'', i.e. for any fixed u

$$\mathbb{E}[\exp\{j\, u\, Z''_N\}] \longrightarrow \exp\{-\tfrac{1}{2}\, u^2\, V''\}$$

then the r.v. $Z_N = Z'_N + Z''_N$ converges in distribution to a Gaussian r.v. with zero mean and variance $V = V' + V''$, as $N \uparrow \infty$

▶ initialization : recall that $X_0^{i,f} \sim \eta_0$ and $\bar{\mu}_0^f = \eta_0$, hence

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N [\phi(X_0^{i,f}) - \int_{\mathbb{R}^m} \phi(x)\, \bar{\mu}_0^f(dx)\,] \Longrightarrow \mathcal{N}(0, \bar{v}_0^f(\phi))$$

in distribution as $N \uparrow \infty$, with asymptotic variance

$$\bar{v}_0^f(\phi) = \mathrm{var}(\phi, \eta_0) = \int_{\mathbb{R}^m} |\phi(x)|^2\, \eta_0(dx) - |\int_{\mathbb{R}^m} \phi(x)\, \eta_0(dx)\,|^2$$

▶ forecast step : recall that $\bar{\mu}_k^f = \bar{\mu}_{k-1}^a \, T_k$, where

$$T_k \, \phi(x) = \int_{\mathbb{R}^m} \phi(f_k(x) + w) \; p_k^W(dw)$$

Proposition asymptotic variance of forecast approximation

$$\bar{v}_k^f(\phi) = \bar{v}_{k-1}^a(T_k \, \phi) + \sigma_k^{2,f}(\phi)$$

in terms of

- asymptotic variance of analysis approximation at previous step, evaluated for a transformed test function
- asymptotic Monte Carlo variance

$$\sigma_k^{2,f}(\phi) = \int_{\mathbb{R}^m} T_k \, |\phi|^2(x) \; \bar{\mu}_{k-1}^a(dx) - \int_{\mathbb{R}^m} |T_k \, \phi(x)|^2 \; \bar{\mu}_{k-1}^a(dx)$$

hint :

$$
\begin{aligned}
Z_N &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\phi(X_k^{i,f}) - \int_{\mathbb{R}^m} \phi(x')\, \bar{\mu}_k^f(dx')\,] \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\phi(f_k(X_{k-1}^{i,a}) + W_k^i) - \int_{\mathbb{R}^m} \phi(x')\, \bar{\mu}_k^f(dx')\,] \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\,\phi(f_k(X_{k-1}^{i,a}) + W_k^i) - T_k\, \phi(X_{k-1}^{i,a})\,] \\
&\qquad\qquad + \frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\,T_k\, \phi(X_{k-1}^{i,a}) - \int_{\mathbb{R}^m} T_k\, \phi(x)\, \bar{\mu}_{k-1}^a(dx)\,] \\
&= Z'_N + Z''_N
\end{aligned}
$$

▶ analysis step : recall that $\bar{\mu}_k^a = \bar{\mu}_k^f \, T_k^{\mathrm{KF}}(\bar{\mu}_k^f)$, where

$$T_k^{\mathrm{KF}}(\bar{\mu}_k^f) \, \phi(x) = \int_{\mathbb{R}^d} \phi(x + K_k(\bar{P}_k^f) \, (Y_k - H_k \, x - v)) \, q_k^V(v) \, dv$$

Proposition asymptotic variance of analysis approximation

$$\bar{v}_k^a(\phi) = \bar{v}_k^f(Q_k^{\mathrm{KF}}(\bar{\mu}_k^f) \, \phi) + \sigma_k^{2,a}(\phi)$$

in terms of

- asymptotic variance of analysis approximation at previous step, evaluated for a transformed test function

- asymptotic Monte Carlo variance

$$\sigma_k^{2,a}(\phi) = \int_{\mathbb{R}^m} T_k^{\mathrm{KF}}(\bar{\mu}_k^f) \, |\phi|^2(x) \, \bar{\mu}_k^f(dx) - \int_{\mathbb{R}^m} |T_k^{\mathrm{KF}}(\bar{\mu}_k^f) \, \phi(x)|^2 \, \bar{\mu}_k^f(dx)$$

here, new transform

$$Q_k^{\mathrm{KF}}(\bar{\mu}_k^f)\,\phi(x) = T_k^{\mathrm{KF}}(\bar{\mu}_k^f)\,\phi(x) + (x - \bar{m}_k^f)^*\,M_k^{\mathrm{KF}}(\bar{\mu}_k^f,\phi)\,(x - \bar{m}_k^f)$$

is defined in terms of matrices

$$M_k^{\mathrm{KF}}(\bar{\mu}_k^f,\phi) = H_k^*\,(H_k\,\bar{P}_k^f\,H_k^* + R_k)^{-1}\,L_k^{\mathrm{KF}}(\bar{\mu}_k^f,\phi)\,(I - K_k(\bar{P}_k^f)\,H_k)$$

and

$$L_k^{\mathrm{KF}}(\bar{\mu}_k^f,\phi) \;=\; \int_{\mathbb{R}^m}\int_{\mathbb{R}^d}(Y_k - H_k\,x - v)\;\phi'(x + K_k(\bar{P}_k^f)\,(Y_k - H_k\,x - v))$$

$$q_k^V(v)\,dv\;\bar{\mu}_k^f(dx)$$

hint :

$$
\begin{aligned}
Z_N &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\phi(X_k^{i,a}) - \int_{\mathbb{R}^m} \phi(x')\; \bar{\mu}_k^a(dx')\,] \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\phi(X_k^{i,f} + K_k(P_k^{N,f})\; (Y_k - H_k\, X_k^{i,f} - V_k^i)) - \int_{\mathbb{R}^m} \phi(x')\; \bar{\mu}_k^a(dx')\,] \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\,\phi(X_k^{i,f} + K_k(P_k^{N,f})\; (Y_k - H_k\, X_k^{i,f} - V_k^i)) - T_k^{\mathrm{KF}}(\mu_k^{N,f})\,\phi(X_k^{i,f})\,] \\
&\qquad + \frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\,T_k^{\mathrm{KF}}(\mu_k^{N,f})\,\phi(X_k^{i,f}) - T_k^{\mathrm{KF}}(\bar{\mu}_k^f)\,\phi(X_k^{i,f})\,] \\
&\qquad + \frac{1}{\sqrt{N}} \sum_{i=1}^{N} [\,T_k^{\mathrm{KF}}(\bar{\mu}_k^f)\,\phi(X_k^{i,f}) - \int_{\mathbb{R}^m} T_k^{\mathrm{KF}}(\bar{\mu}_k^f)\,\phi(x)\; \bar{\mu}_k^f(dx)\,] \\
&= Z'_N + Z''_N + Z'''_N
\end{aligned}
$$

▶ practical computations : iterating the recurrence relations

$$\bar{v}_k^a(\phi) = \bar{v}_k^f(Q_k^{\mathrm{KF}}(\bar{\mu}_k^f)\,\phi) + \sigma_k^{2,a}(\phi)$$

and

$$\bar{v}_k^f(\phi) = \bar{v}_{k-1}^a(T_k\,\phi) + \sigma_k^{2,f}(\phi)$$

yields

$$\begin{aligned}
\bar{v}_k^a(\phi) &= \bar{v}_k^f(Q_k^{\mathrm{KF}}(\bar{\mu}_k^f)\,\phi) + \sigma_k^{2,a}(\phi) \\
&= \bar{v}_{k-1}^a(\underbrace{T_k\,Q_k^{\mathrm{KF}}(\bar{\mu}_k^f)}_{R_k^{\mathrm{KF}}(\bar{\mu}_k^f)}\,\phi) + \underbrace{\sigma_k^{2,f}(Q_k^{\mathrm{KF}}(\bar{\mu}_k^f)\,\phi) + \sigma_k^{2,a}(\phi)}_{\sigma_k^2(\phi)}
\end{aligned}$$

with initialization

$$\bar{v}_0^a(\phi) = \bar{v}_0^f(Q_0^{\mathrm{KF}}(\bar{\mu}_0^f)\,\phi) + \sigma_0^{2,a}(\phi) = \mathrm{var}(Q_0^{\mathrm{KF}}(\eta_0)\,\phi, \eta_0) + \sigma_0^{2,a}(\phi)$$

writing $R_k^{\mathrm{KF}} = R_k^{\mathrm{KF}}(\bar{\mu}_k^f)$ for simplicity

$$\begin{aligned}
\bar{v}_k^a(\phi) &= \bar{v}_{k-1}^a(R_k^{\mathrm{KF}}\,\phi) + \sigma_k^2(\phi) \\
\bar{v}_{k-1}^a(R_k^{\mathrm{KF}}\,\phi) &= \bar{v}_{k-2}^a(R_{k-1}^{\mathrm{KF}}\,R_k^{\mathrm{KF}}\,\phi) + \sigma_{k-1}^2(R_k^{\mathrm{KF}}\,\phi) \\
&\vdots \\
\bar{v}_l^a(R_{l+1}^{\mathrm{KF}}\cdots R_k^{\mathrm{KF}}\,\phi) &= \bar{v}_{l-1}^a(R_l^{\mathrm{KF}}\cdots R_k^{\mathrm{KF}}\,\phi) + \sigma_l^2(R_{l+1}^{\mathrm{KF}}\cdots R_k^{\mathrm{KF}}\,\phi) \\
&\vdots \\
\bar{v}_1^a(R_2^{\mathrm{KF}}\cdots R_k^{\mathrm{KF}}\,\phi) &= \bar{v}_0^a(R_1^{\mathrm{KF}}\cdots R_k^{\mathrm{KF}}\,\phi) + \sigma_1^2(R_2^{\mathrm{KF}}\cdots R_k^{\mathrm{KF}}\,\phi)
\end{aligned}$$

hence

$$\bar{v}_k^a(\phi) = \bar{v}_0^a(R_1^{\mathrm{KF}}\cdots R_k^{\mathrm{KF}}\,\phi) + \sum_{l=1}^{k} \sigma_l^2(R_{l+1}^{\mathrm{KF}}\cdots R_k^{\mathrm{KF}}\,\phi)$$

in terms of backward–propagated functions

$$R_{l+1:k}^{\mathrm{KF}}\,\phi = R_{l+1}^{\mathrm{KF}}\cdots R_k^{\mathrm{KF}}\,\phi$$

further simplifications occur in the special case of

- linear (and quadratic) test functions ϕ
- linear drift function f_k

indeed

- forward–propagated distributions $\bar{\mu}_k^f$ and $\bar{\mu}_k^a$ are Gaussian distributions with moments given by the Kalman filter
- backward–propagated functions $R_{l+1}^{\mathrm{KF}} \cdots R_k^{\mathrm{KF}}\, \phi$ remain quadratic at all steps

and explicit calculations can be obtained