Some issues and results on the EnKF and particle filters for meteorological models

Chaos 2009 Conference

Christophe Baehr & Olivier Pannekoucke

Météo-France / IMT-LSP Univ. Paul Sabatier

Χανιά, Κρήτη

Friday, June 05th, 2009

METEO FRANCE
Toujours un temps d'avance

What shall we talk about ?

1. The nonlinear filtering problem and Particle Filter resolution.
2. The EnKF convergence and mean-field interpretation.
3. The number of particles in a PF. What about the dimension ?
4. The coupling of Local PF and EnKF, a first solution ?
5. The Pointwise PF, another way to face the dimensionality.
6. Further developpements.

The nonlinear filtering problem
Particle Filter resolution

Nonlinear State System

- The State System fixes the filtering problem.

$$\begin{cases} X_{n+1} &= X_n + A_n(X_n)\Delta t + B_n(X_n)\Delta W_n \\ Y_n &= H_n(X_n) + \sigma V_n \end{cases}$$

where V_n and W_n are Wiener process or not. A_n, B_n are $\mathbb{R}^d$-valued functions of X_n and H_n a $\mathbb{R}^{d'}$-valued function.

- The associated filtering problem is to compute $\mathcal{L}aw(X_{[0,n]}|\mathcal{Y}_n)$

- The filtering estimator and the prediction estimator are defined by :

$$\begin{aligned} \hat{\eta}_n(f) &= \mathbb{E}[f(X_0,\ldots,X_n) \mid Y_0 = y_0, \ldots Y_n = y_n] \\ \eta_n(f) &= \mathbb{E}[f(X_0,\ldots,X_n) \mid Y_0 = y_0, \ldots Y_{n-1} = y_{n-1}] \end{aligned}$$

Nonlinear Filtering

Nonlinear State System

- With the Filtering State System we define :
  - the Markovian transition kernel M_n associated with

$$X_{n+1} = X_n + A_n(X_n)\Delta t + B_n(X_n)W_n$$

  - A (non-unique) selection kernel S_{n,η_n} using :
    $G_n(x_n) \stackrel{def}{=} g_n(x_n, y_n)$ with
    $\mathbb{P}(H(x_n) + \sigma V_n \in dy_n \mid X_n = x_n) \; = \; g_n(x_n, y_n)\lambda_n(dy_n)$
- The sequential nonlinear filtering algorithm is :

$$\eta_n \xrightarrow{S_{n,\eta_n}} \hat{\eta}_n = \eta_n \; S_{n,\eta_n} \xrightarrow{M_{n+1}} \eta_{n+1} = \hat{\eta}_n \; M_{n+1}$$

The selection kernel is non-unique

- The most classical, and historically the first described :

$$S_{n,\eta_n}(x_n, dx) = \frac{G_n(x)}{\eta_n(G_n)} \eta_n(dx)$$

- The kernel with genetic selection :

$$S_{n,\eta_n}(x_n, dx) = G_n(x_n)\delta_{x_n}(dx) + [1 - G_n(x_n)] \frac{G_n(x)}{\eta_n(G_n)} \eta_n(dx)$$

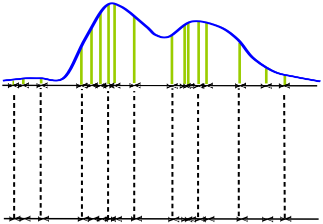
- The genetic selection with a parameter ε_n :

$$S_{n,\eta_n}(x_n, dx) = \varepsilon_n \, . \, G_n(x_n)\delta_{x_n}(dx) + [1 - \varepsilon_n \, . \, G_n(x_n)] \frac{G_n(x)}{\eta_n(G_n)} \eta_n(dx)$$

where $\varepsilon_n \, . \, G_n(x_n) \in [0, 1]$ for any $x_n \in E$

Particle approximation of the nonlinear filtering problem

- These integral equations are not analytically computable.
- We may use an approximation method to solve these equations.
- We use a particle resolution for the probability laws, and here to solve the filtering problem.
- This particle filtering belongs to SMC methods.



Particle approximation of the nonlinear filtering problem

- Beginning the *nth* time step, we start with the set of particles $(\hat{\xi}_n^{i,N})_{1\leq i\leq N}$.

- $(\hat{\xi}_n^{1,N},\ldots,\hat{\xi}_n^{N,N})$ are distributed according to

$$\lim_{N\to\infty}\sum_{i=1}^{N}\delta_{\hat{\xi}_n^{i,N}} = \mathcal{L}aw(X_n \mid Y_0,\ldots,Y_n)$$

the nth marginal of $\mathcal{L}aw(X_0,\ldots,X_n \mid Y_0,\ldots,Y_n)$

- The population of particles with its N offsprings moves with the selection/mutation algorithm .
  - The prediction (mutation) is following the dynamics model.
  - In a genetic case, the selection keeps alive the accepted states and redistributes only the rejected particles.

The nonlinear filtering problem
The Ensemble Kalman Filter method

The Ensemble Kalman Filter

- The EnKF is an empirical estimator.
- The EnKF tries to reproduce the dynamics of the classical Kalman Filter.
- The EnKF uses several realizations of a nonlinear model to estimate empirically the covariance errors matrices.
- The EnKF is convergent but not tends to the Optimal Filter as we will see.

EnKF Algorithm

- The point of departure : we give $P_0 > 0$ and an ensemble of points $(\hat{X}_0^1, \ldots \hat{X}_0^N)$ i.i.d. with $\mathcal{N}(0, P_0)$
- For any $i = 1 \ldots N$, and $n \geq 1$ we have a 1st step :

$$
\begin{aligned}
X_n^i &= A_n(\hat{X}_{n-1}^i) + \sqrt{Q_n}.W_n^i && W_n^i \sim \mathcal{N}(0,1) \\
m_n^N &= \frac{1}{N}\sum_{i=1}^N X_n^i \\
P_n^N &= \frac{1}{N-1}\sum_{i=1}^N (X_n^i - m_n^N)(X_n^i - m_n^N)^T \\
R_n^N &= \frac{1}{N}\sum_{i=1}^N \sqrt{R_n}.V_n^i.(\sqrt{R_n}.V_n^i)^T && V_n^i \sim \mathcal{N}(0,1) \\
G_n^N &= P_n^N.C_n^T.(C_n.P_n^N.C_n^T + R_n^N)^{-1}
\end{aligned}
$$

- Then a correction step :

$$\hat{X}_n^i = X_n^i + G_n^N.[Y_n - C_n.X_n^i + \sqrt{R_n}.V_n^i]$$

Mean-field process interpretation of the EnKF

Mean-field process interpretation

- We combine the previous equations to have :

$$\hat{X}_n^i = A_n(\hat{X}_{n-1}^i) + \sqrt{Q_n}.W_n^i + G_n^N.\left[Y_n - C_n.A_n(\hat{X}_{n-1}^i) - C_n.\sqrt{Q_n}.W_n^i + \sqrt{R_n}.V_n^i\right]$$

- The EnKF approaches the Markovian process Z_n (and not the filter process)

$$Z_n = A_n(Z_{n-1}) + \sqrt{Q_n}.W_n + G_n.\left[Y_n - C_n.A_n(Z_{n-1}) - C_n.\sqrt{Q_n}.W_n + \sqrt{R_n}.V_n\right]$$

with $G_n = P_n \; C_n^T [C_n \; P_n \; C_n^T \; + \; R_n]^{-1}$ and
$P_n = \mathbb{E}\big([A_n(Z_{n-1}) + \sqrt{Q_n} \; W_n][A_n(Z_{n-1}) + \sqrt{Q_n} \; W_n]^T\big)$,
$R_n = \mathbb{E}\big([\sqrt{R_n} \; V_n]\sqrt{R_n} \; V_n^T\big)$

Summary : EnKF vs PF

- A Particle Filter converges to the Optimal Filter as the number of particle goes to infinity.

- A EnKF converges a process Z_n as the number of elements goes to infinity.

- The estimation of an EnKF is optimal if the pair (X_n, Y_n) is linear Gaussian, and in the other cases the EnKF is only the best linear estimator.

- For an equal number of elements, the Particle Filter is cheaper than the EnKF.

- For a small ensemble, a Particle Filter is more risky than an EnKF.

- In high dimensional problems, EnKF has no troubles while Particle Filters ask questions.

- There are many Particles Filters, one for each selection rule.

Particle Filters Regimes

Particle Filters and high dimensions

- The (exact) nonlinear filter ($\eta_{n+1}(f) = \eta_n \; K_{n+1,\eta_n}(f)$) is not concerned by the problem of the dimensionality.

- This is the particle approximation which brings the dimension into question.

- C. Snyder et al. suggest a numerical experiment to evaluate the performance of a Particle Filter as the dimension of the state space increases.

Particle Filters and high dimensions

- C. Snyder et al. propose to use the Lorenz-96 model and an filtering algorithm with a classical selection.

- The Lorenz-96 model is a cyclic dynamical system coupling together the dimensions :

$$\frac{dx_\alpha}{dt} = -x_{\alpha-2}.x_{\alpha-1} + x_{\alpha-1}.x_{\alpha+1} - x_\alpha + F$$

where x_α is the αth component of the state vector $X = (x_1, \ldots, x_\alpha, \ldots, x_d)$ and $1 \leq \alpha \leq d$.

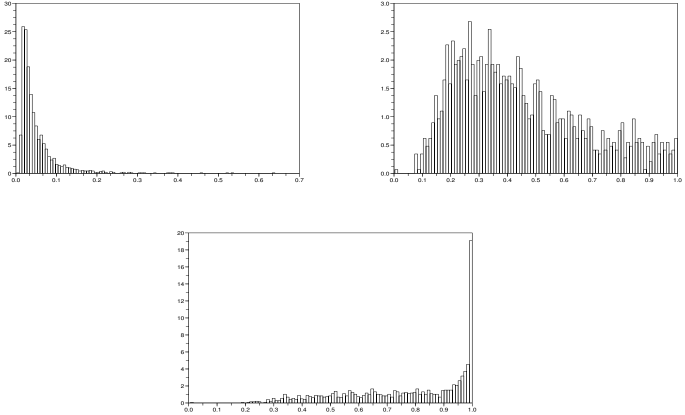
- We propose an experiment where half of the dimensions are directly observed, the others are blind.

- The behaviour of the PF is evaluated by W_n the maximum of particles weight for each time step :

$$W_n = \max_{i \in [1,N]} w_n^i = \max_{i \in [1,N]} \frac{G_n^i}{\sum_{j=1}^N G_n^j}$$

Particle Filters and high dimensions

- For $N = 1000$ particles they obtain the histograms of the maximum weight when the dimensions are $d = 10$, 30 and 80



Genetic Particle Filter faces the dimension

- So it seems to have problems with their Particle Filter when the dimension is (not so !) high.

- We propose to use a genetic selection with a parameter $\varepsilon_n = 1/sup(G_n)$:

$$S_{n,\eta_n}(x_n, dx) = \varepsilon_n \ . \ G_n(x_n)\delta_{x_n}(dx) + [1 - \varepsilon_n \ . \ G_n(x_n)] \frac{G_n(x)}{\eta_n(G_n)} \eta_n(dx)$$

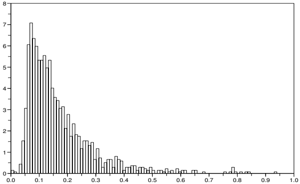
- We introduce a second change in the algorithm : we perturb each particle with a gaussian noise and take it into account in the potential function G_n.

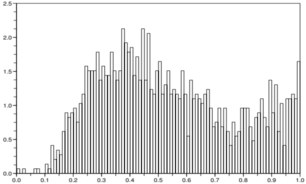
- This changes drastically the behaviour of the particle approximation !

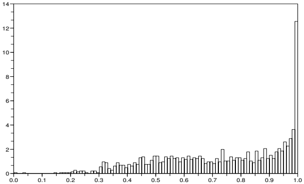
- We have a convergent filter for dimensions higher than 400 for 1000 particles !

Genetic Particle Filter faces the dimension

⊝ For $N = 1000$ particles and dimensions $d = 200$, 600 and 1500 we have the histograms :

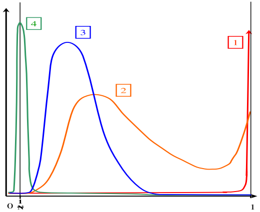






Genetic Particle Filter faces the dimension

- The critical number of particles seems to be $\mathcal{O}(d)$, while *Snyder et al* have found an exponential one.



- Is there a possible theorem ?
  - Case 1 : $\exists\ N_1^{crit}$ s.t. $\forall N < N_1^{crit}$, $\mathbb{P}(W_{max} = 1) = 1$
  - Case 3 : $\exists\ N_2^{crit}$ s.t. $\forall N > N_2^{crit}$, $\mathbb{P}(W_{max} \leq \beta) \geq \alpha$
  - Case 4 : When $\mathbb{P}(|W_{max} - \frac{1}{N}| \leq \beta) = 1$, the algorithm is non-adapted.

The use of Local Particle Filter

Coupling of an EnKF and PF

- Even if the EnKF converges to a process which is not the filtering process, it requires few elements to work in high dimensions.

- Even if the PF converges to the right filtering process, it requires a number of particles in $\mathcal{O}(d)$ for a genetic selection.

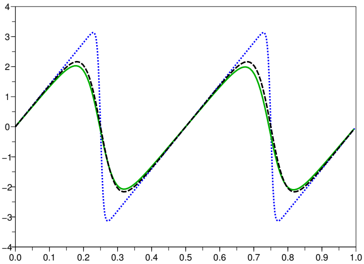
- We propose now to couple together a EnKF and a Local Particle Filter. (I pass over Rao-Blackwellized Particle filters in silence ... may be later for questions if you wish ...)

- For the numerical experiment, we use a dicretized 1D-Burgers equation on $[0, 1]$ with 361 points for an estimation of the true reality and 161 points for the model used by the EnKF.

- Then we use a local model on the first front (161 dimensions on the intervalle $[0, \frac{1}{2}]$) with a Particle Filter to assimilate the observations.

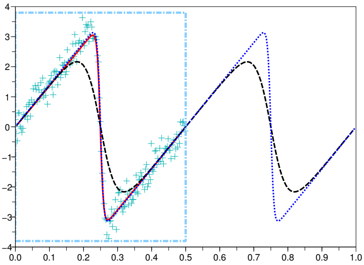
Coupling Local PF and EnKF

EnKF result with 100 elements and after a cycle of 37 assimilations



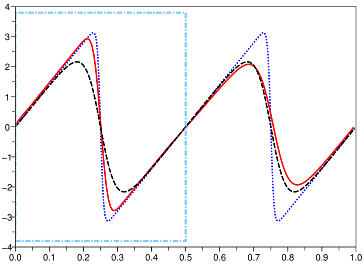
Coupling Local PF and EnKF

Limited-Area Model, with a Local PF using 100 particles after 37 cycles of assimilation.



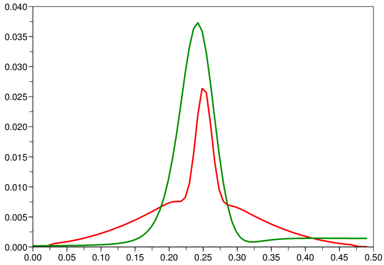
Coupling Local PF and EnKF

Feedback of the particles to the EnKF elements in the LAM domain by randomization.



Coupling Local PF and EnKF

Variance error of the EnKF and the LPF coupled with the EnKF.



The filtering of pointwise turbulent measurements

Pointwise Particle Filter

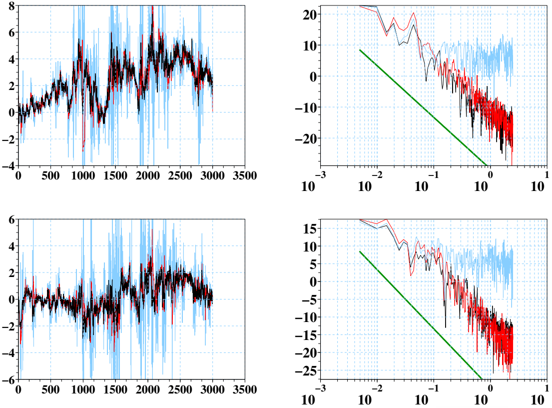
- We can go one step forward with the definition of a Pointwise Particle Filter.
- It requires a stochastic Lagrangian representation of the random medium.
- We have defined the process conditioned to live into a ball centered on each grid point of the model.
- We have developed a Particle Filter for mean-field processes and with a conditioning of the process to the observations.
- This Pointwise Particle Filter need about 100 to 500 particles per grid points.
- This work is in progress for meteorological or turbulent fields.

Example of filtering for turbulent fluid observation

- This is the illustration of one Pointwise Particle Filter for turbulence measurements.
- The model is a 3D stochastic Lagrangian representation of turbulence.
- There is a specific treatement of the Eulerian averages and unobserved parameters.
- We perturb reference observations of real atmospheric wind and temperature, we filter the corrupted signals and we compare the results with the references.
- In this experiment, the Pointwise Particle Filter uses 300 particles.

Example of filtering for turbulent fluid observation

➭ Time series and Power Density Spectra : Toulouse, France, the 18th of July 2006 between 16h58 and 17h00 UTC.



Conclusion

Outcomes and further developpements

- EnKF converges to a mean-field process which is not the optimal filter.
- PF with adapted genetic selection could have an $\mathcal{O}(d)$ critical number of particles.
- PF and EnKF can be coupled together to solve some high dimensional problems.
- Pointwise PF could be a solution for any high dimensional problems.
- We will try to assimilate observation on a 3D turbulent field with PPF or coupled EnKF and LPF for other meteorological models.
- There are other selection kernels for PF to find and study (mixing correction and genetic selection, piloting/tuning of the selection parameter, ...).

Some references

- P. Del Moral. *Feynman-Kac Formulae, Genealogical and Interacting Particle Systems with Applications.* Springer-Verlag, 2004.
- Francois Le Gland, Valérie Monbet, and Vu-Duc Tran. *Large sample asymptotics for the ensemble Kalman filter.* In Dan O. Crisan and Boris L. Rozovskii, editors, Handbook on Nonlinear Filtering. Oxford University Press, to appear.
- C. Snyder, T. Bengtsson, P. Bickel, and J. Anderson. *Obstacles to high-dimensional particle filtering*, Mon. Wea. Rev., 136 :4629-4640, 2008.
- C. Baehr. *Nonlinear Filtering for Observations on a Random Vector Field along a Random Path. Application to Atmospheric Turbulent Velocities*, Submitted to M2AN - ESAIM

- We have 2 schemes for nonlinear filters.

- One for the Particle Filter.

- One for the EnKF.

For any bounded measurable function f

$$\begin{aligned}\eta_{n+1} &= \eta_n \, K_{n+1,\eta_n,\pi_n}(f) = \eta_n \, S_{n,\eta_n} M_{n+1,\pi_n}(f) \\ \eta_{n+1} &= \eta_n \, K_{n+1,\eta_n,\pi_n}(f) = \eta_n \, \mathcal{C}_{n,\eta_n} M_{n+1,\pi_n}(f)\end{aligned}$$