EnKF-based particle filters

Jana de Wiljes, Sebastian Reich, Wilhelm Stannat, Walter Acevedo

June 20, 2017

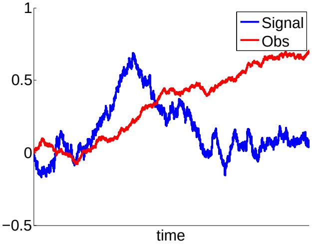
Filtering Problem

Signal

$$\mathrm{d}X_t = f(X_t)\mathrm{d}t + \sqrt{2}C\mathrm{d}W_t$$

Observations

$$dY_t = h(X_t)dt + R^{1/2}dV_t.$$



Goal: determine

$$\pi(x|Y_{0:t})$$

State of the art

Linear Model: Kalman- Bucy Filter

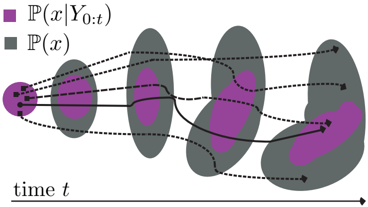
$$\mathrm{d}\bar{x}_t^M = A\bar{x}_t^M \mathrm{d}t + b\mathrm{d}t - P_t^M H^{\mathrm{T}} R^{-1} (H\bar{x}_t^M \mathrm{d}t - \mathrm{d}Y_t)$$

Non-linear Model: approximate with empirical measure, i.e.,

$$\pi(x|Y_{0:t}) \approx \frac{1}{M} \sum_{i=1}^{M} \delta(x - X_t^i).$$

Ansatz: define modified evolution equation for particles X_t^i

Ensemble Kalman Filter (EnKF)



$$\mathrm{d}X_t^i = f(X_t^i)\mathrm{d}t + C\mathrm{d}W_t^i - \frac{1}{2}P_t^M H^{\mathrm{T}} R^{-1}\left(HX_t^i\mathrm{d}t + H\bar{x}_t^M\mathrm{d}t - 2\mathrm{d}Y_t\right)$$

$$\bar{x}_t^M = \frac{1}{M}\sum_{i=1}^{M} X_t^i \qquad P_t^M = \frac{1}{M-1}\sum_{i=1}^{M}(X_t^i - \bar{x}_t^M)(X_t^i - \bar{x}_t^M)^{\mathrm{T}}$$

Works remarkably well in practice: meteorology, oil reservoir exploration

But: theoretical understanding is largely missing

Recent accuracy results for EnKF

Interacting particle representation of the model error:

$$\mathrm{d}X_t^i = f(X_t^i)\mathrm{d}t + CC^\top (P_t^M)^{-1}(X_t^i - \bar{x}_t^M)\mathrm{d}t - \frac{1}{2} P_t^M H^\mathrm{T} R^{-1} \left(HX_t^i \mathrm{d}t + H\bar{x}_t^M \mathrm{d}t - 2\mathrm{d}Y_t \right)$$

Setting: $\mathrm{d}Y_t = X_t \mathrm{d}t + R^{1/2}\mathrm{d}W_t$ with $R = \varepsilon I$

Results: ([dWRS16])

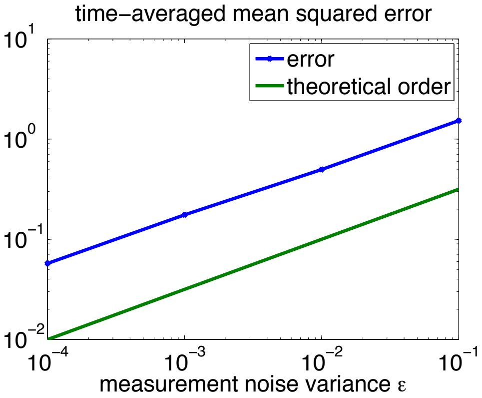
- Control of spectrum of covariance matrix P_t^M over $t \in [0, \infty)$.
- Control of estimation error $e_t = \|X_t^{\mathrm{ref}} - \bar{x}_t^M\|^2$ in expectation

$$\mathbb{E}[E_t] = \mathcal{O}(\epsilon^{1/2}) \tag{1}$$

and pathwise over fixed interval $t \in [0, T]$.

$$\mathbb{P}[\sup_{t \leq T} E_t \geq c_1 \epsilon^q] \leq \mathcal{O}(\epsilon^{1/2-\eta-q}) \tag{2}$$

Numerical confirmation for L63



EnKF-based particle filters

Ansatz: modified evolution equation for the particles e.g., of the form

$$\mathrm{d}X_t^i = f(X_t^i)\mathrm{d}t + C\mathrm{d}W_t^i - \sum_j X_t^j ds_t^{ji}$$

Aims:

- achieve first/second order accuracy
- to go beyond Gaussian assumption
- consistency for $M \to \infty$
- hybrid formulation to combine different interacting particle filters ([CRR16])

Linear ensemble transform filters (LETF)

Given: M samples $x_i^{\mathrm{f}} \sim \pi(x)$ (prior)

Analysis step:

$$x_j^a = \sum_i x_i^f d_{ij} \tag{3}$$

with transformation matrix $D = \{d_{ij}\}$ subject to

$$\sum_{i=1}^{M} d_{ij} = 1 \tag{4}$$

Examples: ENKF, ESRF, NETF, ETPF (see [RC15])

Ensemble transform Particle Filter (ETPF)

Given:

- M samples $x_i^f \sim \pi(x)$ (prior ensemble)
- normalized importance weights $w_i \propto \pi(y|x_i^f)$ (likelihood)

Desired: M samples $x_i^a \sim \pi(x|y)$
Ansatz: replace resampling step with linear transformation by interpreting it as discrete Markov chain given by transition matrix

$$D \in \mathbb{R}^{M \times M}$$

s.t. $d_{ij} \geq 0$ and

$$\sum_i d_{ij} = 1, \qquad \frac{1}{M} \sum_j d_{ij} = w_i.$$

Benefits: localization, hybrid

Ensemble transform Particle Filter (ETPF)

Then the posterior ensemble members are distributed according to the columns of the transformation

$$\tilde{X}_j^a \sim \begin{pmatrix} d_{1j} \\ d_{2j} \\ \vdots \\ d_{Mj} \end{pmatrix} \text{ and } x_j^a = \mathbb{E}[\tilde{X}_j^a] = \sum_i x_i^a d_{ij} \tag{5}$$

Example. Monomial resampling

$$D_{Mono} := \mathbf{w} \otimes \mathbf{1} = \begin{pmatrix} w_1 & w_1 & \cdots & w_1 \\ w_2 & w_2 & \cdots & w_2 \\ \vdots & \vdots & \ddots & \vdots \\ w_M & w_M & \cdots & w_M \end{pmatrix}.$$

Ensemble transform Particle Filter (ETPF)

Then the posterior ensemble members are distributed according to the columns of the transformation

$$\tilde{X}_j^a \sim \begin{pmatrix} d_{1j} \\ d_{2j} \\ \vdots \\ d_{Mj} \end{pmatrix} \text{ and } x_j^a = \mathbb{E}[\tilde{X}_j^a] = \sum_i x_i^a d_{ij} \tag{5}$$

Example. Monomial resampling

$$D_{Mono} := \mathbf{w} \otimes \mathbf{1} = \begin{pmatrix} w_1 & w_1 & \cdots & w_1 \\ w_2 & w_2 & \cdots & w_2 \\ \vdots & \vdots & \ddots & \vdots \\ w_M & w_M & \cdots & w_M \end{pmatrix}.$$

Here: X^f and X^a are independent.
Idea: increase correlation between X^f and X^a

Ensemble transform Particle Filter (ETPF)

Solve optimization problem

$$D_{ETPF} = \arg\min M \sum_{i,j=1}^{M} t_{ij} ||x_i^f - z_j^f||^2$$

to find transformation matrix D_{ETPF} that increases correlation ([Rei13]).

Remarks:

- consistent for $M \to \infty$
- first-order accurate for finite M, i.e,

$$\overline{x}^a = \frac{1}{M} \sum_{i=1}^{M} x_i^a = \sum_{i=1}^{M} w_i x_i^f \qquad (6)$$

- **But**: not second-order accurate

Second-order accuracy

The analysis covariance matrix

$$\widehat{P}^a = \frac{1}{M}\sum_{i=1}^{M}(x_i^a - \overline{x}^a)(x_i^a - \overline{x}^a)^{\mathrm{T}}$$

is equal to the covariance matrix defined by the importance weights, i.e.

$$P^a = \sum_{i=1}^{M} w_i(t_k)(x_i^f - \overline{x}^a)(x_i^f - \overline{x}^a)^T.$$

First-order accurate LETFs

Notation:

$$\mathbf{X}^f = (x_1^f, \dots, x_M^f),\; \mathbf{X}^a = (x_1^a, \dots, x_M^a) \in \mathbb{R}^{N_x \times M}$$

and analogously

$$\mathbf{w} = (w_1, \dots, w_M)^{\mathrm{T}} \in \mathbb{R}^{M \times 1} \text{ and } \mathbf{W} = \mathrm{diag}\,(\mathbf{w})$$

LETF is first-order accurate if

$$\frac{1}{M}\mathbf{X}^{\mathrm{a}}\mathbf{1} = \mathbf{X}^{\mathrm{f}}\mathbf{w}.$$

This holds if D satisfies

$$\frac{1}{M}D\mathbf{1} = \mathbf{w}.$$

First-order accurate LETFs

Class of first-order accurate LETFs

$$\mathcal{D}_1 = \{ D \in \mathbb{R}^{M\times M} |\, D^{\mathrm{T}}\mathbf{1} = \mathbf{1},\ D\mathbf{1} = M\mathbf{w} \}$$

Examples:

- $D_{\mathrm{EnKF}} \notin \mathcal{D}_1$
- $D_{\mathrm{ESRF}} \notin \mathcal{D}_1$
- $D_{\mathrm{ETPF}} \in \mathcal{D}_1$
- $D_{\mathrm{Mono}} \in \mathcal{D}_1$

Second-order accurate LETF

Analysis covariance matrix:

$$\widehat{\mathbf{P}}^{\mathrm{a}} = \frac{1}{M}\mathbf{X}^{\mathrm{f}}(D - \mathbf{w}\mathbf{1}^{\mathrm{T}})(D - \mathbf{w}\mathbf{1}^{\mathrm{T}})^{\mathrm{T}}(\mathbf{X}^{\mathrm{f}})^{\mathrm{T}} \tag{7}$$

Importance sampling covariance matrix:

$$\mathbf{P}^{\mathrm{a}} = \mathbf{X}^{\mathrm{f}}(\mathbf{W} - \mathbf{w}\mathbf{w}^{\mathrm{T}})(\mathbf{X}^{\mathrm{f}})^{\mathrm{T}}. \tag{8}$$

Thus the following equation has to hold for second-order accuracy:

$$(D - \mathbf{w}\mathbf{1}^{\mathrm{T}})(D - \mathbf{w}\mathbf{1}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{W} - \mathbf{w}\mathbf{w}^{\mathrm{T}} \tag{9}$$

Class of second-order accurate LETFs

$$\mathcal{D}_2 = \{D \in \mathcal{D}_1 |\, (D - \mathbf{w}\mathbf{1}^{\mathrm{T}})(D - \mathbf{w}\mathbf{1}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{W} - \mathbf{w}\mathbf{w}^{\mathrm{T}}\}. \tag{10}$$

Second-order corrected LETF

Given: $D \in \mathcal{D}_1$

Goal: correct transformation $\widehat{D} \in \mathcal{D}_2$

Ansatz:

$$\widehat{D} = D + \Delta$$

with $\Delta \in \mathbb{R}^{M \times M}$ such that $\Delta\mathbf{1} = \mathbf{0}$, $\Delta^{\mathrm{T}}\mathbf{1} = \mathbf{0}$ ([dWAR17]).

Need to solve algebraic Riccati equation:

$$M(\mathbf{W} - \mathbf{w}\mathbf{w}^{\mathrm{T}}) - (D - \mathbf{w}\mathbf{1}^{\mathrm{T}})(D - \mathbf{w}\mathbf{1}^{\mathrm{T}})^{\mathrm{T}} \\ = (D - \mathbf{w}\mathbf{1}^{\mathrm{T}})\Delta^{\mathrm{T}} + \Delta(D - \mathbf{w}\mathbf{1}^{\mathrm{T}})^{\mathrm{T}} + \Delta\Delta^{\mathrm{T}}.$$

Numerical example I

Gaussian prior, non-Gaussian likelihood:

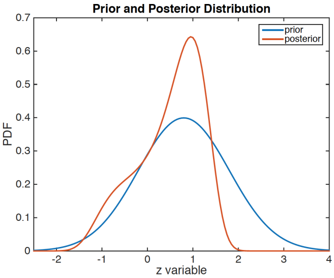


Figure: Prior and posterior distribution for the single Bayesian inference step

Numerical example I

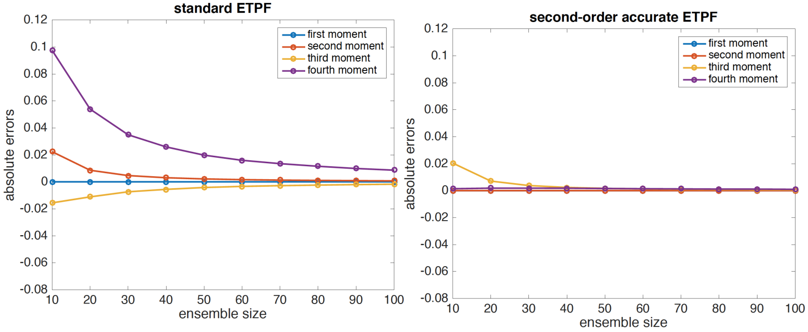


Figure: Absolute errors in the first four moments of the posterior distribution.

Numerical example II

Lorenz-63 model, first component observed infrequently ($\Delta t = 0.12$) and with large measurement noise ($R = 8$):

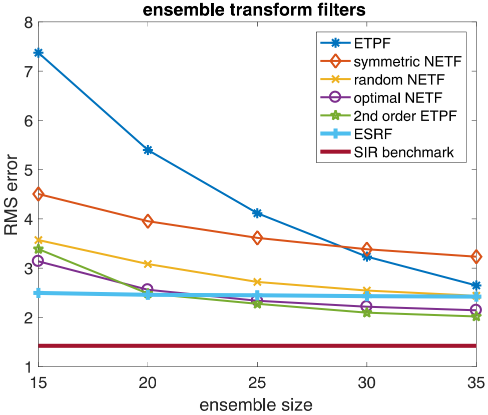


Figure: RMSEs for various second-order accurate LETFs compared to the ETPF, the ESRF, and the SIR PF as a function of the sample size, M.

Numerical example II

Hybrid filter: $D := D_{\rm ESRF}\ D_{\rm ETPF}$.

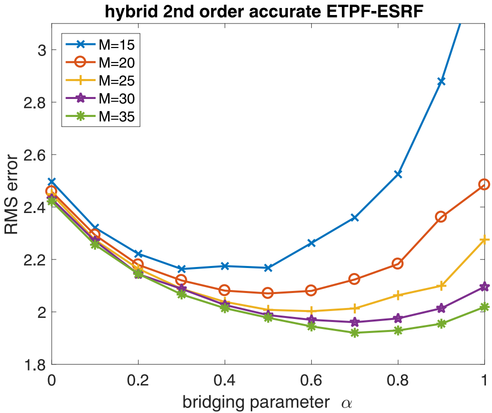


Figure: RMSEs for hybrid ESRF ($\alpha = 0$) and 2nd-order corrected LETF/ETPF ($\alpha = 1$) as a function of the sample size, M.

Data Assimilation Center Potsdam

SFB 1294
Data Assimilation

- PhD and Postdoc positions in 11 projects
- Mathematical foundation and applications in geophysics, biologie and cognitive sciences
- Support for external projects and long term visitors

Data Assimilation Center Potsdam

SFB 1294
Data Assimilation

- PhD and Postdoc positions in 11 projects
- Mathematical foundation and applications in geophysics, biologie and cognitive sciences
- Support for external projects and long term visitors

Thank you!

References I

Nawinda Chustagulprom, Sebastian Reich, and Maria Reinhardt.
A hybrid ensemble transform filter for nonlinear and spatially extended dynamical systems.
SIAM/ASA J Uncertainty Quantification, 4:592–608, 2016.

J. de Wiljes, W. Acevedo, and S. Reich.
A second-order accurate ensemble transform particle filter.
accepted in SIAM Journal on Scientific Computing, (ArXiv:1608.08179), 2017.

J. de Wiljes, S. Reich, and W. Stannat.
Long-time stability and accuracy of the ensemble Kalman-Bucy filter for fully observed processes and small measurement noise.
Technical Report https://arxiv.org/abs/1612.06065, 2016.

S. Reich and C. Cotter.
Probabilistic Forecasting and Bayesian Data Assimilation.
Cambridge University Press, 2015.

References II

Sebastian Reich.

A non-parametric ensemble transform method for Bayesian inference.

SIAM J Sci Comput, 35:A2013–A2014, 2013.