State and parameter estimation from observed signal increments

Nikolas Nüsken[1] Sebastian Reich[1] **Paul J. Rozdeba**[1]

[1]Institute of Mathematics
University of Potsdam, Potsdam, Germany

Joint Workshop on Conservation Principles, Data, and Uncertainty in Atmosphere-Ocean Modelling
University of Potsdam
April 3, 2019

SFB 1294 Data Assimilation

CRC 1114

Deutsche Forschungsgemeinschaft
German Research Foundation

Overview

1. Bayesian data assimilation via particle-based filtering: for model variables q and data y,

$$\pi_t(q|y) \simeq \sum_{i=1}^{M} \pi_t(q^i|y)\,\delta(q - q^i)$$

2. Continuous-time filtering with feedback control.

$$\mathrm{d}X_t = f(X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t + \mathrm{d}U_t(X_t, Y_t)$$

3. Assimilation of SDE state increments $\rightarrow$ filters with correlated noise.
4. Simultaneous state and drift parameter estimation.

Vanilla particle filter

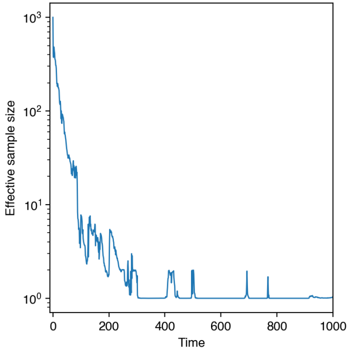
$$\textit{Signal:}\ \mathrm{d}X_t = f(X_t, a)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$$
$$\textit{Data:}\ \mathrm{d}Y_t = h(X_t, a)\,\mathrm{d}t + \rho\,\mathrm{d}V_t$$
$$\pi_t(x|y) = \pi(x(t)|y(:t)) = ?$$

- Particle approximation:

$$\pi_t(q|y) \simeq \sum_{i=1}^{M} \pi_t(q^i|y)\,\delta(q - q^i)$$

- Particle filter is "passive":

$$X_n^i \leftarrow X_{n-1}^i + \sigma\sqrt{\Delta t}\,\xi_n^i$$
$$w_n^i \leftarrow \frac{w_{n-1}^i\,\pi(Y_n|X_n^i)}{\sum_i w_{n-1}^i\,\pi(Y_n|X_n^i)}$$



Vanilla particle filter: state tracking example

$$
\begin{aligned}
\textit{Model:}\ \mathrm{d}X_{1,t} &= X_{2,t}\,\mathrm{d}t + \sigma\,\mathrm{d}W_{1,t} \\
\mathrm{d}X_{2,t} &= (-X_{1,t} - 0.1X_{2,t})\,\mathrm{d}t + \sigma\,\mathrm{d}W_{2,t} \\
\textit{Data:}\ \mathrm{d}Y_t &= X_{1,t}\,\mathrm{d}t + \rho\,\mathrm{d}V_t
\end{aligned}
$$

Particle evolution (pf.mov)

Proposed filter: Key features

1. Filtering in continuous time

$$\textit{Signal:}\ \mathrm{d}X_t = f(X_t, a)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$$
$$\textit{Data:}\ \mathrm{d}Y_t = h(X_t, a)\,\mathrm{d}t + \rho\,\mathrm{d}V_t$$
$$\pi_t(x|y) = \pi(x(t)|y(:t)) = ?$$

2. Particle/ensemble approximation

$$\pi_t(x|y) \simeq \sum_{i=1}^{M} \pi_t(x^i|y)\,\delta(x - x^i)$$

3. Feedback control

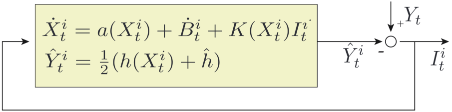


Figure: T. Yang, et al (2013)

State tracking example

$$\begin{aligned} \textit{Model:}\ \mathrm{d}X_{1,t} &= X_{2,t}\,\mathrm{d}t + \sigma\,\mathrm{d}W_{1,t} \\ \mathrm{d}X_{2,t} &= (-X_{1,t} - 0.1X_{2,t})\,\mathrm{d}t + \sigma\,\mathrm{d}W_{2,t} \\ \textit{Data:}\ \mathrm{d}Y_t &= X_{1,t}\,\mathrm{d}t + \rho\,\mathrm{d}V_t \end{aligned}$$

Filter w/ feedback control (enkf.mov)

Key features

1. *Later:* Noisy (partial) observations of SDE increments

$$\mathrm{d}Y_t = H\,\mathrm{d}X_t + \rho\,\mathrm{d}V_t,\ H \in \mathbb{R}^{L\times D_x}$$

2. Extended phase space: model states $x \in \mathbb{R}^{D_x}$ and parameters $a \in \mathbb{R}^{D_a}$

$$z = \{x, a\},\ z \in \mathbb{R}^{D=D_x+D_a}$$

3. $\rho > 0$: simultaneous state and parameter estimation

$$\pi_t(z|y) \simeq \sum_{i=1}^{M} \pi_t(z^i|y)\,\delta(z - z^i) =?$$

4. *Consistency* with true posterior for $M \to \infty$

Feedback particle filter (T. Yang, et al, 2013)

- Uses feedback control idea:

$$\begin{aligned} \mathrm{d}X_t^i &= f(X_t^i, a)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t^i + K_t(X_t^i) \circ \mathrm{d}I_t^i(X_t^i, Y_t) \\ \mathrm{d}Y_t &= h(X_t^*)\,\mathrm{d}t + \rho\,\mathrm{d}V_t \end{aligned}$$

- $M \to \infty$ limit: consistent with particle filter?
- Variational problem: minimize $\mathrm{KL}\left(\pi_t \| \pi_t^*\right)$.
- Gain K solves the elliptic PDE (Euler-Lagrange eq. for min KL)

$$\partial_i \left[\pi_t \left(K_t^{ij}(x)\rho^{jk}\right)\right] = -\pi_t \left(h_t^k(x) - \pi_t[h_t^k(x)]\right) \text{ s.t. } \lim_{x \to \pm\infty} \pi_t K_t(x) = 0$$

- Innovation term $\mathrm{d}I_t^i$ given by

$$\mathrm{d}I_t^i = \mathrm{d}Y_t - \frac{1}{2}\left[h(X_t^i) + \pi_t[h(X_t^i)]\right]\mathrm{d}t$$

$$\text{or } \mathrm{d}I_t^i = \mathrm{d}Y_t - \left[h(X_t^i)\,\mathrm{d}t + \rho\,\mathrm{d}U_t^i\right]$$

Assimilation of SDE increments

- Consider drift functions f of the form:

$$f(x,t) = f_0(x,t) + \sum_{i=1}^{D_a} a_i \varphi_i(x) = f_0(x,t) + \Phi(x)^\mathsf{T} a$$

  e.g. Φ are finite-element basis functions (used in nonparametric estimation).

- Extended phase space/increment measurements:

$$\begin{aligned} \mathrm{d}X_t &= f(X_t, A_t, t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t \\ \mathrm{d}A_t &= 0 \\ \mathrm{d}Y_t &= H\,\mathrm{d}X_t + \rho\,\mathrm{d}V_t \end{aligned}$$

Zero-noise case → ensemble Kalman-Bucy filter (EnKBF)

- Special case: $\rho = 0$, $H = I$, and $\pi_0(x) = \delta(x - X_0)$ leads to

$$\mathrm{d}Y_t = \mathrm{d}X_t = H\left[f(X_t, A_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t\right] \quad (Y_t = X_t)$$

- Now A_t has dynamics from feedback term:

$$\begin{aligned} \mathrm{d}A_t^i &= K(A_t^i) \circ \mathrm{d}I_t^i(A_t, Y_t) \\ \mathrm{d}I_t^i &= \mathrm{d}Y_t - \left[f(Y_t, A_t^i)\,\mathrm{d}t + \sigma\,\mathrm{d}U_t^i\right] \end{aligned}$$

- With π_0 Gaussian, make constant-gain approximation:

$$K_t = P_t^{aa}\Phi(Y_t)^{\mathsf{T}}\sigma^{-2},\ P_t^{aa} \simeq \frac{1}{M-1}\sum_{i=1}^{M}\left[A_t^i - \bar{A}^M\right]\left[A_t^i - \bar{A}^M\right]^{\mathsf{T}}$$

- Nonliner observation function h:

$$K_t = P_t^{ah}\sigma^{-2},\ P_t^{ah} \simeq \frac{1}{M-1}\sum_{i=1}^{M}\left[A_t^i - \bar{A}^M\right]\left[h(A_t^i) - \bar{h}^M\right]^{\mathsf{T}}$$

General case: $\rho > 0$, $H \neq I$

- Unless $H\sigma = 0$, model and measurement errors are correlated:

$$E_t^m = \sigma W_t, \; E_t^o = H\sigma W_t + \rho V_t$$

- Define: $C = H\sigma^2 H^T + \rho^2$
- Modified feedback particle filter equations[1]:

$$\begin{aligned}
\mathrm{d}X_t =& f(X_t, A_t)\,\mathrm{d}t \\
& + \sigma\,\mathrm{d}W_t \left(\nabla_x \Psi_t(X_t, A_t) + \sigma^2 H^T\right) C^{-1} \circ \mathrm{d}I_t + \Omega(X_t) \\
\mathrm{d}A_t =& \nabla_a \Psi_t(X_t, A_t) C^{-1} \circ \mathrm{d}I_t \\
\nabla_z \cdot (\pi_t \nabla_z \Psi_t) =& -\pi_t \left(Hf(z) - \pi_t[Hf(z)]\right)
\end{aligned}$$

[1] N. Nüsken, S. Reich, P.J.R. (2019)

General case → EnKBF

- Gaussian assumptions:

$$\nabla_x \Psi_t = P_t^{xh}, \quad \nabla_a \Psi_t = P_t^{ah},$$
$$P_t^{xh} \simeq \frac{1}{M-1} \sum_{i=1}^{M} \left[X_t^i - \bar{X}_t^M\right] \left[f(X_t^i, A_t^i) - \bar{f}_t^M\right]^\mathsf{T} H^\mathsf{T}$$
$$P_t^{ah} \simeq \frac{1}{M-1} \sum_{i=1}^{M} \left[A_t^i - \bar{A}_t^M\right] \left[f(X_t^i, A_t^i) - \bar{f}_t^M\right]^\mathsf{T} H^\mathsf{T}$$

- Results in EnKF formulation for joint estimation problem

$$\mathrm{d}X_t^i = f(X_t^i, A_t^i)\mathrm{d}t + \sigma\, \mathrm{d}W_t^i + \left(P_t^{xh} + \sigma^2 H^\mathsf{T}\right) C^{-1} \mathrm{d}I_t^i$$
$$\mathrm{d}A_t^i = P_t^{ah} C^{-1} \mathrm{d}I_t^i$$
$$\mathrm{d}I_t^i = \mathrm{d}Y_t - \left[H\left(f(X_t^i, A_t^i)\mathrm{d}t + \sigma \mathrm{d}W_t^i\right) + \rho \mathrm{d}U_t^i\right]$$

Example: Reduced multiscale system[2]

- Linear stochastic system with three time scales:

$$\mathrm{d}X_t = \left(\frac{\sqrt{\sigma}}{\varepsilon} Z_t - \theta X_t\right)\mathrm{d}t$$

$$\mathrm{d}Z_t = -\frac{1}{\varepsilon^2} Z_t\,\mathrm{d}t + \frac{\sqrt{2}}{\varepsilon}\,\mathrm{d}W_t \quad (W_t \sim N(0,t))$$

- Homogenization limit (model reduction):

$$\lim_{\varepsilon\to 0} \to \mathrm{d}\widetilde{X}_t = -a\widetilde{X}_t\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}\widetilde{W}_t \quad (\widetilde{W}_t \sim N(0,t))$$

- DA: estimate effective parameter a from data process

$$\mathrm{d}Y_t = \mathrm{d}X_t + \rho\,\mathrm{d}V_t \text{ (from \textit{full} system)}$$

- Theory predicts $a = \theta$.

[2] G. Pavliotis and A. Stuart (2008)

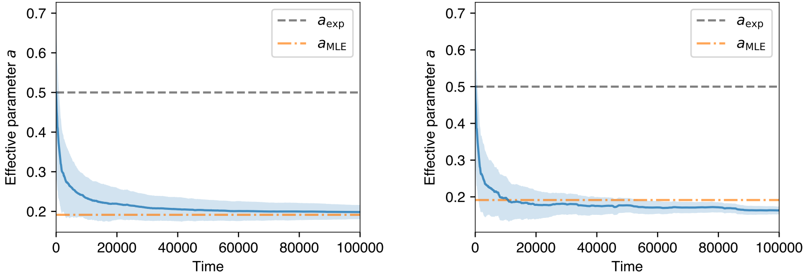
Example: Reduced multiscale system

- Standard likelihood estimation for a via Girsanov's formula:

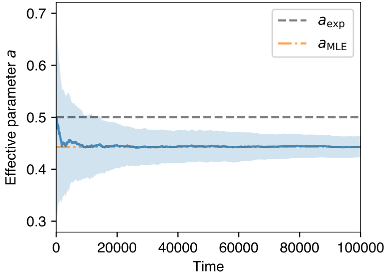
$$\widehat{a}_{\mathrm{ML}} = \frac{\int \widetilde{X}_t \, \mathrm{d}\widetilde{X}_t}{\int |\widetilde{X}_t|^2 \, \mathrm{d}t} \simeq \frac{\sum_{n=1}^{N-1} \widetilde{X}_n(\widetilde{X}_{n+1} - \widetilde{X}_n)}{|\widetilde{X}_n|^2 \Delta t}$$

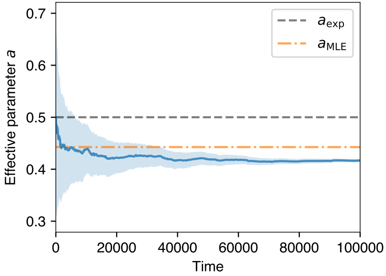
- Inconsistent continuous-time limit: as $\Delta t \to 0$, $\widehat{a}_{\mathrm{ML}} \to 0$.
- Fixed with subsampling: sample solutions with finite Δt
- Data assimilation experiment:
  - Generate reference solution from full system
  - Sample increment data (with noise) from x component
  - Estimate a in reduced model with and without subsampling
  - Prior $\pi_0 = \delta(x - X_0)\, N_a(0, \sigma_0^a)$
  - Draw $M = 100$ particles from prior, run filter with various time steps Δt

Without subsampling: $\Delta t = 0.01$



With subsampling: $\Delta t = 0.1$





Example II: Nonparametric drift estimation

- Fully observed one-dimensional system with model error:

$$\begin{aligned} \mathrm{d}X_t &= f_0(X_t, t)\,\mathrm{d}t + \lambda f_1(X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t \\ \mathrm{d}Y_t &= \mathrm{d}X_t + \rho\,\mathrm{d}V_t \end{aligned}$$

- Filtering problem: $\pi_t(f_1|y) = ?$
- Gaussian process (GP) prior for f_1:

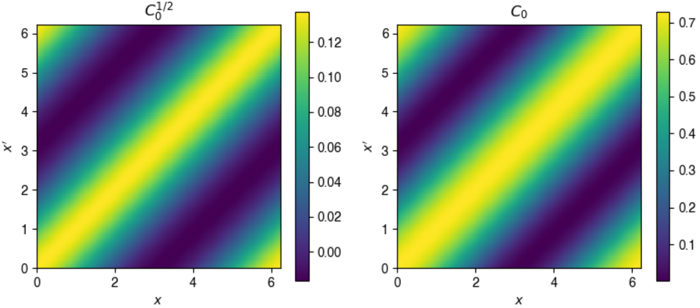
$$\pi_0(f_1(x)) = \mathcal{GP}\left[\mu_0(x), \Gamma_0(x, x')\right]$$
$$\mu_0 = 0,\ \Gamma_0 \equiv \mathcal{Q}_0^{-1} = \left\{\eta\left[(-\Delta)^p + \kappa\mathcal{I}\right]\right\}^{-1}$$

- Consider finite-dimensional decomposition:

$$f_1(x) = \sum_{j=1}^{D_a} a_j \varphi_j(x)$$

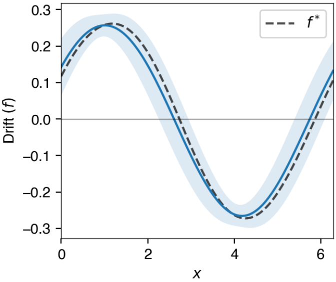
for D_a functions φ_j with compact support (i.e. tent functions).

Nonparametric drift estimation: Prior covariance



Example II: Nonparametric drift estimation

- Numerical experiment:
  - Draw ground truth f_1^* from prior GP with known η, p, κ.
  - Generate reference solution with known X_0, sample increment data.
  - Prior $\pi_0 = \delta(x - X_0)\, GP_f(0, C_0)$ (GP is discretized in space)



Summary and ongoing work

- Derived feedback particle filter formulation for simultaneous state and parameter estimation with noisy increment measurements.
- Robust approximation in the form of ensemble Kalman-Bucy filter.
- Further extensions to multi-scale (reduced) models?
- Nonparametric case: posterior convergence & consistency; (on-line) prior selection.

References

1. A. Bain & D. Crisan, "Fundamentals of Stochastic Filtering". Springer-Verlag (2009).
2. N. Nusken, S. Reich, and P.J.R., "State and parameter estimation from observed signal increments" (submitted). arXiv:1903.10717, 2019.
3. G. Pavliotis and A. Stuart, "Multiscale Methods". Springer-Verlag (2008).
4. T. Yang, P.G. Mehta, and S.P. Meyn, "Feedback Particle Filter". IEEE Trans. Aut. Control **58(10)**, 2013.

Thank you for your attention!

Contact: paul.rozdeba@uni-potsdam.de