Bayesian filters for continuous-discrete dynamic models

Joaquín Miguez

Department of Signal Theory & Communications
Universidad Carlos III de Madrid

STOCHASTICA Directions Workshop
Cork, 29 April 2026.

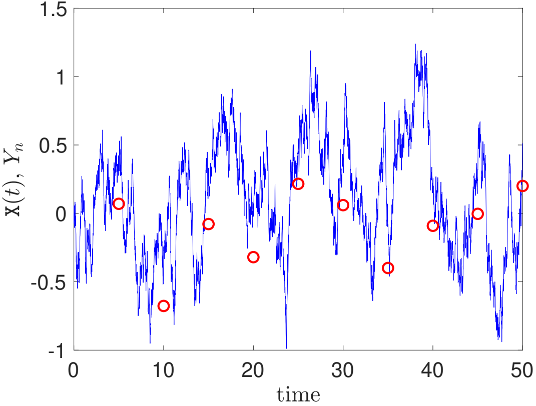
.
.
.

Partially supported by grant PID2024-158181NB-I00 NISA, funded by MCIN/AEI/10.13039/501100011033 and ERDF, and grant IDEA-CM TEC-2024/COM-89, funded by Community of Madrid. Part of this research was performed at the mathematical research institute MATRIX.

Contents

Continuous-discrete state space model

- Continuous-time stochastic state $\mathtt{X}(t)$ (——)
- Discrete-time observations $Y_n = m(\mathtt{X}(t_n)) + \text{noise}$ ($\circ$)



- The aim is to compute $\mathbb{P}\left(\mathtt{X}(t) \in \mathrm{d}x \mid \{Y_n, n \in \mathbb{N} : t_n \leq t\}\right)$.

Continuous-discrete state space models

- Assume the signal of interest is an Itô process $\mathtt{X}(t)$ in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The dynamics are described by the SDE

$$\mathrm{d}\mathtt{X}(t) = \overbrace{a(\mathtt{X}(t), t)}^{\text{drift}}\,\mathrm{d}t + \underbrace{\sigma(\mathtt{X}(t), t)}_{\text{diffusion coeff.}}\overbrace{\mathrm{d}\mathtt{W}(t)}^{\text{Wiener proc.}} \quad , \quad t \in [0, T].$$

- For a given time grid $0 = t_0 < t_1 < \cdots < t_n < \cdots < t_N = T$:
  - Let $X_n = \mathtt{X}(t_n)$, the observation at time t_n is

$$Y_n = m(X_n) + U_n, \quad n = 1, 2, \ldots$$

    where U_n is iid noise and $m : \mathcal{X} \mapsto \mathcal{Y}$ is an observation function.
  - The sequence $X_n := \mathtt{X}(t_n)$, $n \geq 0$, is described by the Markov kernels

$$K_n(x', \mathrm{d}x) := \mathbb{P}\left(X_n \in \mathrm{d}x | X_{n-1} = x'\right), \quad \text{for } n \geq 1.$$

- The likelihood at time t_n, denoted $g_n(x) \propto \bar{g}(y_n - m(x))$, depends on the law $\bar{g}(u)$ of the noise variables U_n and the observed value $Y_n = y_n$.
  Example: if $U_n \sim \mathcal{N}(0, \Sigma_u) \Longrightarrow g_n(x) \propto \exp\left\{-\frac{1}{2}\|y_n - m(x)\|^2_{\Sigma_u^{-1}}\right\}$

Bayesian filtering

- Given the kernels K_n and likelihoods g_n, we are interested in the marginal probability laws

$$\overbrace{\xi_n(\mathrm{d}x) := \mathbb{P}(X_n \in \mathrm{d}x | Y_{1:n-1} = y_{1:n-1})}^{\text{one-step-ahead prediction}}, \quad \underbrace{\pi_n(\mathrm{d}x) := \mathbb{P}(X_n \in \mathrm{d}x | Y_{1:n} = y_{1:n})}_{\text{Bayesian filter}}$$

- Recursive construction

$$\begin{aligned} \text{(predict)} \quad \xi_n(\mathrm{d}x) &= K_n \pi_{n-1}(\mathrm{d}x) := \int_{\mathcal{X}} K_n(x', \mathrm{d}x) \pi_{n-1}(\mathrm{d}x'), \\ \text{(update)} \quad \pi_n(\mathrm{d}x) &= \frac{g_n(x)\xi_n(\mathrm{d}x)}{\xi_n(g_n)}, \quad \text{where } \xi_n(g_n) = \int_{\mathcal{X}} g_n(x)\xi_n(\mathrm{d}x) \end{aligned}$$

- Issues: in general...
  - we can only sample K_n (approximately),
  - both integrals are intractable.

Need for efficient numerical methods!

Approximation errors

Approximation pipeline

✓ Continuous-discrete model

$\Downarrow$

time discretisation of SDE

$\Downarrow$

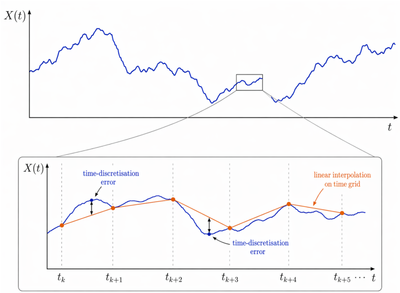
≈ Discrete-time model

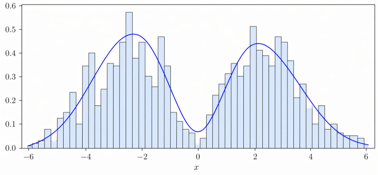
$\Downarrow$

approx. numerical algorithm

$\Downarrow$

≈ Sequence of filtering distributions





Time-discretisation errors

- The kernel K_n is intractable. We can (approximately) sample, e.g.,

$$
\begin{aligned}
\bar{X}^h_{n,0} &= x', \quad \text{and} \\
\bar{X}^h_{n,j} &= \bar{X}^h_{n,j-1} + ha(\bar{X}^h_{n,j-1}, t'_{n,j-1}) + \sqrt{h}\sigma(\bar{X}^h_{n,j-1}, t'_{n,j-1})Z_{n,j},
\end{aligned}
$$

  for $j = 1, \dots, K$ and $h = t'_{n,j} - t'_{n,j-1}$. Then $X^h_n := \bar{X}^h_{n,J} \approx X_n = \mathtt{X}(t_n)$.
- True vs. approximate model: given observations $Y_n = y_n$,

$$
\overbrace{\begin{array}{rcl} X_0 & \sim & \pi_0(dx) \\ X_n & \sim & K_n(x_{n-1}, dx) \\ g_n(X_n) & \propto & \bar{g}(y_n - m(X_n)) \end{array}}^{\text{continuous time}} \quad , \quad \underbrace{\begin{array}{rcl} X^h_0 & \sim & \pi_0(dx) \\ X^h_n & \sim & K^h_n(x_{n-1}, dx) \\ g_n(X^h_n) & \propto & \bar{g}(y_n - m(X_n)) \end{array}}_{\text{discrete time}}
$$

A choice of numerical scheme yields an approximate kernel K^h. For a numerical scheme of weak order p

$$
\sup_{0 \le n \le N} \left| \mathbb{E}[\phi(X_n)] - \mathbb{E}\left[\phi\left(X^h_n\right)\right] \right| < C_\phi h^p \quad \text{for a test function } \phi \in C^4_B(\mathbb{R}^{d_x}).
$$

Time-discretisation errors

- Time-discretisation affects the posterior marginals

$$\overbrace{\begin{aligned} \xi_n(\mathrm{d}x) &= (K_n \pi_{n-1})(\mathrm{d}x) \\ \pi_n(\mathrm{d}x) &= \frac{g_n(x)\xi_n(\mathrm{d}x)}{\xi_n(g_n)} \end{aligned}}^{\text{continuous time}}, \quad \underbrace{\begin{aligned} \xi_n^h(\mathrm{d}x) &= (K_n^h \pi_{n-1}^h)(\mathrm{d}x) \\ \pi_n^h(\mathrm{d}x) &= \frac{g_n(x)\xi_n^h(\mathrm{d}x)}{\xi_n^h(g_n)} \end{aligned}}_{\text{discrete time}}$$

Theorem (Time-discretisation, Akyildiz *et al.* (2024); Erdogan *et al.* (2026))

Let $Y_{1:N} = y_{1:N}$ be arbitrary but fixed, let K_n^h be a weak-order 1 scheme and choose $\phi \in C_B^4(\mathbb{R}^{d_x})$ with $\|\phi\|_\infty \leq 1$. If

- $g_n > 0$, $\|g_n\|_\infty \leq 1$ *and* $g_n \in C_B^4(\mathbb{R}^{d_x})$, *and*
- $\bar{\phi}_n \in C_B^4(\mathbb{R}^{d_x})$, *where* $\bar{\phi}_n(x') = \int \phi(x) K_n(x', \mathrm{d}x)$,

then there are finite constants $\{\tilde{C}_n, C_n\}_{n=1}^N$, independent of h, such that

$$\left|\xi_n(\phi) - \xi_n^h(\phi)\right| \leq \tilde{C}_n h \quad \text{and} \quad \left|\pi_n(\phi) - \pi_n^h(\phi)\right| \leq C_n h \quad \text{for } n = 1, \ldots, N.$$

Monte Carlo error example: the particle filter

- Numerical filtering methods add approximation error on top of the discretisation error.

Example: standard particle filter

- Algorithm
  1. Draw $x_0^{h,i} \sim \pi_0$, $i = 1, \ldots, M$.
  2. Recursive step: for $i = 1, ..., M$
     i. draw $\bar{x}_n^{h,i} \sim K_n^h(x_{n-1}^{h,i}, \mathrm{d}x)$,
     ii. weight $w_n^i \propto g_n(\bar{x}_n^{h,i})$,
     iii. resample $\{\bar{x}_n^{h,i}, w_n^i\}_{i=1}^M \longrightarrow \{x_n^{h,i}\}_{i=1}^M$
  3. Approximation $\pi_n^{h,M}(\mathrm{d}x) = \frac{1}{M}\sum_{i=1}^M \delta_{x_n^{i,h}}(\mathrm{d}x)$.
- Approximation error (Erdogan *et al.*, 2026): under smoothness conditions, there are constants $C_n, B_n < \infty$ such that

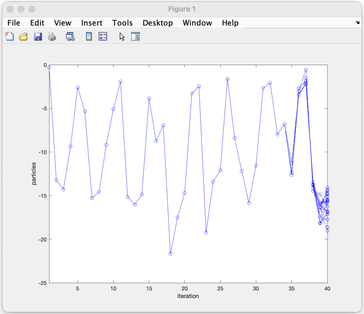
$$\mathbb{E}\left[\left|\pi_n(\phi) - \pi_n^{h,M}(\phi)\right|^p\right]^{\frac{1}{p}} < \overbrace{C_n h}^{\text{discretisation}} + \underbrace{\frac{B_n}{\sqrt{M}}}_{\text{Monte Carlo}}, \quad p \geq 1.$$

Modified SDEs & Monte Carlo filters

Particle filters

Generic particle filter with sampling kernel Q_n & weight degeneracy

1. Draw $x_0^i \sim \pi_0$, $i = 1, \dots, M$.
2. Recursive step: for $i = 1, \dots, M$
   i. draw $\bar{x}_n^i \sim Q_n(x_{n-1}^i, \mathrm{d}x)$,
   ii. weight $w_n^i \propto g_n(\bar{x}_n^i) \frac{\mathrm{d}K_n}{\mathrm{d}Q_n}(x_{n-1}^i, \bar{x}_n^i)$,
   iii. resample $\{\bar{x}_n^i, w_n^i\}_{i=1}^M \longrightarrow \{x_n^i\}_{i=1}^M$



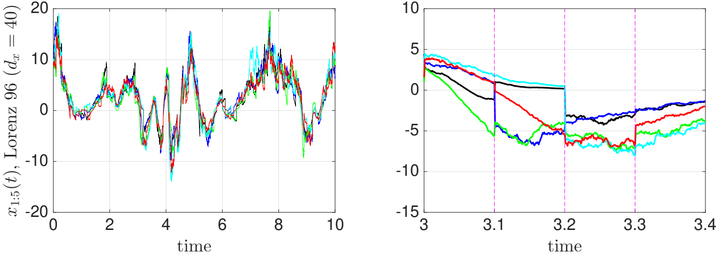
The "optimal" kernel is $Q_n(x_{n-1}^i, \mathrm{d}x) = \mathbb{P}(X_n \in \mathrm{d}x \mid X_{n-1} = x_{n-1}^i, Y_n = y_n)$.

Weight degeneracy is an acute problem in high dimensional models!

https://www.dropbox.com/scl/fi/cxm3v6sqliintpdjq7k9y/path_degeneracy.mov?rlkey=obuzkatt0g30rng8eaeh9p1y9&dl=0

Feedback particle filter

- If weight degeneracy is a problem... remove the weights.
- Prediction using the SDE as in a standard PF.
  Update maps predictive-law particles into filtering-law particles



- Conceptually appealing, but
  - "jumps" in sampled paths,
  - update only approximate (no exact sampling!)

Feedback particle filter (Yang *et al.*, 2014)

- Assume pdfs $\xi_n(x)$ (prediction) and $\pi_n(x)$ (filter).
  How do we convert particle $\bar{x}_n^i \sim \xi_n$ into particle $x_n^i \sim \pi_n$?
- Construct the homotopy of densities

$$\rho_n(x,\tau) \propto \xi_n(x) \mathrm{e}^{\tau\Phi(x)}, \quad \text{where} \quad \Phi(x) = m(x)^\top y_n - \frac{1}{2\sigma^2}\|m(x)\|^2$$

  This yields $\rho_n(x,0) = \xi_n(x)$ and $\rho_n(x,1) \propto \pi_n(x)$, and

$$\partial_\tau \rho = \rho \left[(m(x) - (m,\rho))^\top y_n - \frac{1}{2\sigma^2}\|m(x)\|^2 + \frac{1}{2\sigma^2}(m^2,\rho) \right]$$

- Liouville (continuity) equation:

$$\text{if} \qquad \rho \left[(m(x) - (m,\rho))^\top y_n - \frac{1}{2\sigma^2}\|m(x)\|^2 + \frac{1}{2\sigma^2}(m^2,\rho) \right] = -\nabla \cdot (\rho U)$$

$$\text{then} \qquad \frac{\mathrm{d}\tilde{x}}{\mathrm{d}\tau} = U(\tilde{x},\tau), \quad \text{with} \quad \tilde{x}(0) = \bar{x}_n^i, \quad \text{yields} \quad \tilde{x}(1) = x_n^i \sim \pi_n$$

Finding $U(\tilde{x},\tau)$ requires the numerical solution of a high dimensional Poison equation $\Longrightarrow$ not easy to implement + approximation errors.

Feedback particle filter (Yang *et al.*, 2014)

- Assume pdfs $\xi_n(x)$ (prediction) and $\pi_n(x)$ (filter).
  How do we convert particle $\bar{x}_n^i \sim \xi_n$ into particle $x_n^i \sim \pi_n$?
- Construct the homotopy of densities

$$\rho_n(x,\tau) \propto \xi_n(x) e^{\tau \Phi(x)}, \quad \text{where} \quad \Phi(x) = m(x)^\top y_n - \frac{1}{2\sigma^2}\|m(x)\|^2$$

This yields $\rho_n(x,0) = \xi_n(x)$ and $\rho_n(x,1) \propto \pi_n(x)$, and

$$\partial_\tau \rho = \rho \left[(m(x) - (m,\rho))^\top y_n - \frac{1}{2\sigma^2}\|m(x)\|^2 + \frac{1}{2\sigma^2}(m^2,\rho) \right]$$

- Liouville (continuity) equation:

$$\text{if} \qquad \rho \left[(m(x) - (m,\rho))^\top y_n - \frac{1}{2\sigma^2}\|m(x)\|^2 + \frac{1}{2\sigma^2}(m^2,\rho) \right] = -\nabla \cdot (\rho U)$$

$$\text{then} \qquad \frac{d\tilde{x}}{d\tau} = U(\tilde{x},\tau), \quad \text{with} \quad \tilde{x}(0) = \bar{x}_n^i, \quad \text{yields} \quad \tilde{x}(1) = x_n^i \sim \pi_n$$

Finding $U(\tilde{x},\tau)$ requires the numerical solution of a high dimensional Poison equation $\Longrightarrow$ not easy to implement + approximation errors.

Feedback particle filter (Yang *et al.*, 2014)

- Assume pdfs $\xi_n(x)$ (prediction) and $\pi_n(x)$ (filter).
  How do we convert particle $\bar{x}_n^i \sim \xi_n$ into particle $x_n^i \sim \pi_n$?
- Construct the homotopy of densities

$$\rho_n(x,\tau) \propto \xi_n(x) \mathrm{e}^{\tau \Phi(x)}, \quad \text{where} \quad \Phi(x) = m(x)^\top y_n - \frac{1}{2\sigma^2}\|m(x)\|^2$$

  This yields $\rho_n(x,0) = \xi_n(x)$ and $\rho_n(x,1) \propto \pi_n(x)$, and

$$\partial_\tau \rho = \rho \left[(m(x) - (m,\rho))^\top y_n - \frac{1}{2\sigma^2}\|m(x)\|^2 + \frac{1}{2\sigma^2}(m^2,\rho) \right]$$

- Liouville (continuity) equation:

$$\text{if} \qquad \rho \left[(m(x) - (m,\rho))^\top y_n - \frac{1}{2\sigma^2}\|m(x)\|^2 + \frac{1}{2\sigma^2}(m^2,\rho) \right] = -\nabla \cdot (\rho U)$$

$$\text{then} \qquad \frac{\mathrm{d}\tilde{x}}{\mathrm{d}\tau} = U(\tilde{x},\tau), \quad \text{with} \quad \tilde{x}(0) = \bar{x}_n^i, \quad \text{yields} \quad \tilde{x}(1) = x_n^i \sim \pi_n$$

Finding $U(\tilde{x},\tau)$ requires the numerical solution of a high dimensional Poison equation $\Longrightarrow$ not easy to implement + approximation errors.

The Doob h-transform (Rogers & Williams, 2000)

Recall the particle filter

1. Draw $x_0^i \sim \pi_0$, $i = 1, \dots, M$.
2. Recursive step: for $i = 1, \dots, M$
   i. draw $\bar{x}_n^i \sim Q_n(x_{n-1}^i, \mathrm{d}x)$,
   ii. weight $w_n^i \propto g_n(\bar{x}_n^i) \frac{\mathrm{d}K_n}{\mathrm{d}Q_n}(x_{n-1}^i, \bar{x}_n^i)$,
   iii. resample $\{\bar{x}_n^i, w_n^i\}_{i=1}^M \longrightarrow \{x_n^i\}_{i=1}^M$

"Optimal" $Q_n(x_{n-1}^i, \mathrm{d}x_n) := \mathbb{P}\left(X_n \in \mathrm{d}x \mid X_{n-1} = x_{n-1}^i, Y_n = y_n\right)$.

- How to implement Q_n? Modify the original SDE

$$\mathrm{d}\mathtt{X}(t) = a(\mathtt{X}(t), t)\mathrm{d}t + \sigma(\mathtt{X}(t), t)\mathrm{d}\mathtt{W}(t)$$
$$\Downarrow$$
$$\mathrm{d}\tilde{\mathtt{X}}(t) = \left[a(\tilde{\mathtt{X}}(t), t) + \sigma^2(\tilde{\mathtt{X}}(t), t)\nabla \log h(\tilde{\mathtt{X}}(t), t)\right] \mathrm{d}t + \sigma(\tilde{\mathtt{X}}(t), t)\mathrm{d}\mathtt{W}(t)$$

for $t \in (t_{n-1}, t_n]$, where $\sigma^2(\cdot,\cdot) = \sigma(\cdot,\cdot)\sigma(\cdot,\cdot)^\top$ and

$$h(x, t) := \mathbb{E}\left[g_n\left(\tilde{\mathtt{X}}(t_n)\right) \mid \mathtt{X}(t) = x\right]$$

is a Doob h-transform.

Particle filter via Doob h-transform

SDE for $t \in (t_{n-1}, t_n]$, with initial condition $\mathtt{X}^i(t_{n-1}) = \tilde{\mathtt{X}}^i(t_{n-1}) = x^i_{n-1}$

$$\mathrm{d}\mathtt{X}^i(t) = a(\mathtt{X}^i(t), t)\mathrm{d}t + \sigma(\mathtt{X}^i(t), t)\mathrm{d}\mathtt{W}^i(t)$$
$$\Downarrow$$
$$\mathrm{d}\tilde{\mathtt{X}}^i(t) = \left[a(\tilde{\mathtt{X}}^i(t), t) + \sigma^2(\tilde{\mathtt{X}}^i(t), t)\nabla \log h(\tilde{\mathtt{X}}^i(t), t)\right] \mathrm{d}t + \sigma(\tilde{\mathtt{X}}^i(t), t)\mathrm{d}\mathtt{W}^i(t)$$

- Let $\mathbb{P}^i_t$ and $\tilde{\mathbb{P}}^i_t$ denote the laws of $\mathtt{X}^i(t)$ and $\tilde{\mathtt{X}}^i(t)$. Standard theory shows that, for $t_{n-1} < t \leq t_n$,

$$\frac{\mathrm{d}\mathbb{P}^i_t}{\mathrm{d}\tilde{\mathbb{P}}^i_t}(x) = \frac{h(\mathtt{X}^i(t_{n-1}), t_{n-1})}{h(x, t)}, \quad \text{where} \quad h(x,t) = \mathbb{E}\left[g_n(\mathtt{X}^i(t_n)) \mid \mathtt{X}(t) = x\right]$$

- Hence, at time $t = t_n$, we have

$$\frac{\mathrm{d}\mathbb{P}^i_{t_n}}{\mathrm{d}\tilde{\mathbb{P}}^i_{t_n}}(x) = \frac{\mathbb{E}\left[g_n\left(\mathtt{X}^i(t_n)\right) \mid \mathtt{X}^i(t_{n-1}) = x^i_{n-1}\right]}{g_n(x)}$$

and the factor $g_n(x)$ cancels in the weight eq. $w_n(x) \propto g_n(x)\frac{\mathrm{d}\mathbb{P}^i_t}{\mathrm{d}\tilde{\mathbb{P}}^i_t}(x)$,

$$w^i_n \propto \mathbb{E}\left[g_n\left(\mathtt{X}^i(t_n)\right) \mid \mathtt{X}^i(t_{n-1}) = x^i_{n-1}\right] \quad \text{independently of } \bar{x}^i_n := \tilde{\mathtt{X}}^i(t_n)$$

Particle filter via Doob h-transform

SDE for $t \in (t_{n-1}, t_n]$, with initial condition $\mathtt{X}^i(t_{n-1}) = \tilde{\mathtt{X}}^i(t_{n-1}) = x^i_{n-1}$

$$d\mathtt{X}^i(t) = a(\mathtt{X}^i(t), t)dt + \sigma(\mathtt{X}^i(t), t)d\mathtt{W}^i(t)$$
$$\Downarrow$$
$$d\tilde{\mathtt{X}}^i(t) = \left[a(\tilde{\mathtt{X}}^i(t), t) + \sigma^2(\tilde{\mathtt{X}}^i(t), t)\nabla \log h(\tilde{\mathtt{X}}^i(t), t)\right] dt + \sigma(\tilde{\mathtt{X}}^i(t), t)d\mathtt{W}^i(t)$$

- Let $\mathbb{P}^i_t$ and $\tilde{\mathbb{P}}^i_t$ denote the laws of $\mathtt{X}^i(t)$ and $\tilde{\mathtt{X}}^i(t)$. Standard theory shows that, for $t_{n-1} < t \leq t_n$,

$$\frac{d\mathbb{P}^i_t}{d\tilde{\mathbb{P}}^i_t}(x) = \frac{h(\mathtt{X}^i(t_{n-1}), t_{n-1})}{h(x, t)}, \quad \text{where} \quad h(x, t) = \mathbb{E}\left[g_n(\mathtt{X}^i(t_n)) \mid \mathtt{X}(t) = x\right]$$

- Hence, at time $t = t_n$, we have

$$\frac{d\mathbb{P}^i_{t_n}}{d\tilde{\mathbb{P}}^i_{t_n}}(x) = \frac{\mathbb{E}\left[g_n\left(\mathtt{X}^i(t_n)\right) \mid \mathtt{X}^i(t_{n-1}) = x^i_{n-1}\right]}{g_n(x)}$$

and the factor $g_n(x)$ cancels in the weight eq. $w_n(x) \propto g_n(x)\frac{d\mathbb{P}^i_t}{d\tilde{\mathbb{P}}^i_t}(x)$,

$$w^i_n \propto \mathbb{E}\left[g_n\left(\mathtt{X}^i(t_n)\right) \mid \mathtt{X}^i(t_{n-1}) = x^i_{n-1}\right] \quad \text{independently of } \bar{x}^i_n := \tilde{\mathtt{X}}^i(t_n)$$

Particle filter via Doob h-transform

SDE for $t \in (t_{n-1}, t_n]$, with initial condition $\mathtt{X}^i(t_{n-1}) = \tilde{\mathtt{X}}^i(t_{n-1}) = x^i_{n-1}$

$$\mathrm{d}\mathtt{X}^i(t) = a(\mathtt{X}^i(t), t)\mathrm{d}t + \sigma(\mathtt{X}^i(t), t)\mathrm{d}\mathtt{W}^i(t)$$

$$\Downarrow$$

$$\mathrm{d}\tilde{\mathtt{X}}^i(t) = \left[a(\tilde{\mathtt{X}}^i(t), t) + \sigma^2(\tilde{\mathtt{X}}^i(t), t)\nabla \log h(\tilde{\mathtt{X}}^i(t), t)\right]\mathrm{d}t + \sigma(\tilde{\mathtt{X}}^i(t), t)\mathrm{d}\mathtt{W}^i(t)$$

- Let $\mathbb{P}^i_t$ and $\tilde{\mathbb{P}}^i_t$ denote the laws of $\mathtt{X}^i(t)$ and $\tilde{\mathtt{X}}^i(t)$. Standard theory shows that, for $t_{n-1} < t \le t_n$,

$$\frac{\mathrm{d}\mathbb{P}^i_t}{\mathrm{d}\tilde{\mathbb{P}}^i_t}(x) = \frac{h(\mathtt{X}^i(t_{n-1}), t_{n-1})}{h(x, t)}, \quad \text{where} \quad h(x, t) = \mathbb{E}\left[g_n(\mathtt{X}^i(t_n) \mid \mathtt{X}(t) = x\right]$$

- Hence, at time $t = t_n$, we have

$$\frac{\mathrm{d}\mathbb{P}^i_{t_n}}{\mathrm{d}\tilde{\mathbb{P}}^i_{t_n}}(x) = \frac{\mathbb{E}\left[g_n\left(\mathtt{X}^i(t_n)\right) \mid \mathtt{X}^i(t_{n-1}) = x^i_{n-1}\right]}{g_n(x)}$$

and the factor $g_n(x)$ cancels in the weight eq. $w_n(x) \propto g_n(x)\frac{\mathrm{d}\mathbb{P}^i_t}{\mathrm{d}\tilde{\mathbb{P}}^i_t}(x)$,

$$w^i_n \propto \mathbb{E}\left[g_n\left(\mathtt{X}^i(t_n)\right) \mid \mathtt{X}^i(t_{n-1}) = x^i_{n-1}\right] \quad \text{independently of } \bar{x}^i_n := \tilde{\mathtt{X}}^i(t_n)$$

Particle filter via Doob h-transform: summary

SDE for $t \in (t_{n-1}, t_n]$, with initial condition $\mathtt{X}^i(t_{n-1}) = \tilde{\mathtt{X}}^i(t_{n-1}) = x^i_{n-1}$

$$\mathrm{d}\mathtt{X}^i(t) = a(\mathtt{X}^i(t), t)\mathrm{d}t + \sigma(\mathtt{X}^i(t), t)\mathrm{d}\mathtt{W}^i(t)$$
$$\Downarrow$$
$$\mathrm{d}\tilde{\mathtt{X}}^i(t) = \left[a(\tilde{\mathtt{X}}^i(t), t) + \sigma^2(\tilde{\mathtt{X}}^i(t), t)\nabla \log h(\tilde{\mathtt{X}}^i(t), t)\right]\mathrm{d}t + \sigma(\tilde{\mathtt{X}}^i(t), t)\mathrm{d}\mathtt{W}^i(t)$$

Sampling, optimal kernel Q_n

- Sample $\bar{x}^i_n := \tilde{\mathtt{X}}^i_{t_n}$ with initial condition $\tilde{\mathtt{X}}^i(t_{n-1}) = x^i_{n-1}$
- Then $\bar{x}^i_n \sim \mathbb{P}(\mathrm{d}X_n \mid X_{n-1} = x^i_{n-1}, Y_n = y_n)$, the "optimal" importance function
- The weight $w^i_n \propto \mathbb{E}\left[g_n\left(\mathtt{X}^i(t_n)\right) \mid \mathtt{X}^i(t_{n-1}) = x^i_{n-1}\right]$ is independent of $\bar{x}^i_n$.

Difficulties

Need to compute $\nabla \log h(\tilde{\mathtt{X}}^i(t), t)$ and $\mathbb{E}\left[g_n\left(\mathtt{X}^i(t_n)\right) \mid \mathtt{X}^i(t_{n-1}) = x^i_{n-1}\right]$.
See Chopin *et al.* (2023) for a neural network based implementation and Pieper-Sethmacher *et al.* (2025b,a) for linear approximations.

Constrained filters

- We have assumed so far that $\mathtt{X}(t) \in \mathcal{X} \subseteq \mathbb{R}^{d_x}$. The *truncation* or *constraint* of the state space $\mathcal{X}$ by a sequence of compact subsets $\mathcal{C}_n \subset \mathcal{X}$ at times t_n has proved useful
  - algorithmically, to improve robustness (Garcia-Fernandez *et al.*, 2012; Wang *et al.*, 2020) and
  - theoretically, to improve stability (Crisan *et al.*, 2020; Erdogan *et al.*, 2025).
- A constraint is expressed by a sequence of compact subsets $\mathcal{C} = \{\mathcal{C}_n\}_{n \geq 0}$. The constraint modifies the prior law of X_0 and the Markov kernels

$$\hat{\pi}_0(\mathrm{d}x) = \frac{\mathbb{1}_{\mathcal{C}_0}(x)\pi_0(\mathrm{d}x)}{\pi_0(\mathcal{C}_0)}, \quad \hat{K}_n(x', \mathrm{d}x) = \frac{\mathbb{1}_{\mathcal{C}_n}(x)K_n(x', \mathrm{d}x)}{\int_{\mathcal{X}} \mathbb{1}_{\mathcal{C}_n}(\tilde{x})K_n(x', \mathrm{d}\tilde{x})}, \quad n \geq 1.$$

  As a consequence, we obtain constrained predictions & filters

$$\hat{\xi}_n(\mathrm{d}x) = \hat{K}_n\hat{\pi}_{n-1}(\mathrm{d}x), \quad \hat{\pi}_n(\mathrm{d}x) = g_n(x)\hat{\xi}_n(\mathrm{d}x)/\hat{\xi}_n(g_n), \quad n \geq 1.$$

How do we simulate $\hat{K}_n(x', \mathrm{d}x)$, a constrained SDE?

- Rejection sampling?
- Doob h-transform $h_n(x, t) = \mathbb{P}\left(\mathtt{X}(t_n) \in \mathcal{C}_n \mid \mathtt{X}(t_{n-1}) = x_{n-1}^i\right)$?

High-probability tubes

- Let us use Doob transforms to impose the constraints: for $t \in (t_{n-1}, t_n]$

$$\mathrm{d}\mathtt{X}^n(t) = \left[a(\mathtt{X}^n(t), t) + \sigma^2(\mathtt{X}^n(t), t)\nabla \log h_n(\mathtt{X}^n(t), t)\right] \mathrm{d}t + \sigma(\mathtt{X}^n(t), t)\mathrm{d}\mathtt{W}(t),$$

where $h_n(x, t) := \mathbb{P}\left(\mathtt{X}(t_n) \in \mathcal{C}_n \mid \mathtt{X}(t) = x\right)$.

- Construct a flow of superlevel sets of the Doob h-transform

$$S_n(t) := \{x : h_n(x, t) \geq \exp(-r(t))\}, \quad t_{n-1} < t \leq t_n,$$

generated by a map $r : [t_{n-1}, t_n] \mapsto (0, \infty)$, such that $\lim_{t \to t_n} r(t) = \infty$ to ensure $S_n(t_n) = \mathcal{C}_n$.

- Given $r(t)$, there is some $\epsilon_r \geq 0$ such that

$$\mathbb{P}\left(\mathtt{X}^n(t) \in S_n(t)\right) \geq 1 - \epsilon_r, \quad \forall t \in (t_{n-1}, t_n].$$

If $r(t)$ is selected suitably, $S_n(t)$ captures most of the probability mass of the constrained process $\mathtt{X}^n(t) \implies$ it is a high probability tube connecting $\mathtt{X}(t_{n-1})$ with the terminal set $\mathcal{C}_n$.

How do we implement it?

Barrier functions (Erdogan *et al.*, 2025)

- Approximate high-probability tube
  - Define the constraint by way of superlevel sets of the log-likelihood

$$\mathcal{C}_n := \{x \in \mathcal{X} : \log g_n(x) > \log \zeta_n\}, \quad \text{for a threshold } \zeta_n > 0.$$

  - Define the flow of compact sets

$$\hat{S}_n(t) := \{x \in \mathcal{X} : \log g^n(x,t) > \log \zeta_n\}, \quad t \in [t_{n-1}, t_n],$$

    with interpolated likelihoods

$$g^n(x, t_{n-1}) = g_{n-1}(x) \ \text{ and } \ g^n(x, t_n) = g_n(x).$$

  - $\hat{S}_n(t)$ interpolates $\hat{S}_n(t_{n-1}) = \mathcal{C}_{n-1}$ and $\hat{S}_n(t_n) := \mathcal{C}_n$
- The tube $\hat{S}_n(t)$ is imposed as a soft-constraint,

$$\mathrm{d}\mathtt{X}^n(t) = \left[a(\mathtt{X}^n(t), t) - \mu\sigma^2(\mathtt{X}^n(t), t)\nabla\overbrace{\log b_n(\mathtt{X}^n(t), t)}^{\log-\text{barrier}} \right] \mathrm{d}t + \sigma(\mathtt{X}^n(t), t)\mathrm{d}\mathtt{W}(t)$$

The barrier $-b_n(x,t)$ is a tractable surrogate for the Doob transform $h_n(x,t)$

Implementation of barrier functions

- We have used a shifted soft-plus function with parameters ϱ and $\kappa \geq 1$,

$$u_{\varrho,\kappa}(z) := \frac{1}{\kappa} \log\left(1 + \exp\{\kappa(z-\varrho)\}\right).$$

Parameter ϱ is a radius that determines the boundaries of the tube: for large enough κ,

$$u_{\varrho,\kappa}(z) \approx \begin{cases} 0 & \text{when} \quad z < \varrho \\ z - \varrho & \text{when} \quad z > \varrho \end{cases}$$

- The barrier function is then constructed as

$$\log b_n(x,t) := u_{\varrho,\kappa}(-\log g^n(x,t)), \quad t \in [t_{n-1}, t_n].$$

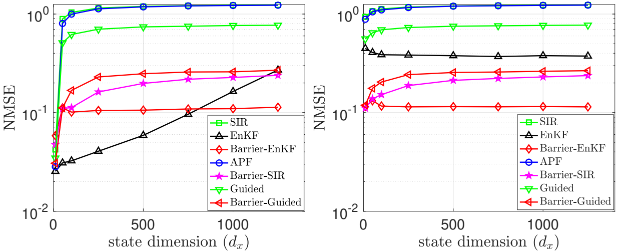
Example: Gaussian observations $Y_n = H^\top X_n + U_n$, $U_n \sim N(0, \sigma_y^2 I)$

The tube is
$\hat{S}_n(t) = \{x \in \mathcal{X} : \log g^n(x,t) \geq \log \zeta_n\} = \left\{x \in \mathcal{X} : \|y^n(t) - H^\top x\| \leq \sqrt{-\frac{\log \zeta}{2\sigma_y^2}}\right\}$ and
$\nabla \log b_n(x,t) = \frac{1}{\sigma_y^2} s\left(\kappa\left(\frac{1}{2\sigma_y^2}\|y^n(t) - H^\top x\|^2 - \varrho_n\right)\right) H(H^\top x - y^n(t))$, where
$\varrho_n = -\log \zeta_n$ and $s(z) = 1/(1+\mathrm{e}^{-z})$ is the logistic function.

Numerics: stochastic Lorenz 96 w/ additive noise

$$\mathtt{d}\mathtt{X}_i(t) = [\mathtt{X}_{i-1}(t)\,(\mathtt{X}_{i+1}(t) - \mathtt{X}_{i-2}(t)) - \mathtt{X}_i(t) + F]\,dt + \sigma_x \mathtt{dW}_i(t), \quad i = 0, \ldots, d_x - 1.$$
$$Y_n = H^\top X_n + U_n, \; U_n \sim N(0, \sigma_y^2 I).$$

Monte Carlo filters with $M = 750$ (left) and $M = \lfloor 0.5 d_x \rfloor$ (right) particles



state dim. d_x	1000	1250	2500	5000	7500	10,000
Barrier-ENKF, $M = 750$	0.1056	0.1129	0.1214	0.1485	0.1604	0.1801
Barrier-SIR, $M = 750$	0.2176	0.2208	0.2560	0.2705	0.2757	0.2794
Barrier-Guided, $M = 750$	0.2478	0.2562	0.2770	0.2762	0.2797	0.2910

Discussion

Discussion

- Bayesian filtering for discretely-observed diffusions (continuous-discrete state space models)
- Time-discretisation & Monte Carlo errors
- Efficient conditional sampling of SDEs
  - Feedback particle filter (Yang *et al.*, 2014)
  - Doob *h*-tranforms (Chopin *et al.*, 2023; Pieper-Sethmacher *et al.*, 2025b,a)

  are hardly practical in high-dimensional, nonlinear settings
- Constrained filters and barrier functions
  - borrowed from control theory (Margellos & Lygeros, 2011; Clark, 2021)
  - barriers are tractable surrogates of intractable Doob *h*-tranforms
  - promising numerics for high-dimensional filtering

Akyildiz, D., Crisan, D., & Miguez, J. 2024. Sequential discretisation schemes for a class of stochastic differential equations and their application to Bayesian filtering. *SIAM Journal on Numerical Analysis*, **62**(2), 946–973.

Chopin, Nicolas, Fulop, Andras, Heng, Jeremy, & Thiery, Alexandre H. 2023. Computational Doob h-transforms for online filtering of discretely observed diffusions. *Pages 5904–5923 of: International Conference on Machine Learning*. PMLR.

Clark, Andrew. 2021. Control barrier functions for stochastic systems. *Automatica*, **130**, 109688.

Crisan, Dan, López-Yela, Alberto, & Miguez, Joaquin. 2020. Stable approximation schemes for optimal filters. *SIAM/ASA Journal on Uncertainty Quantification*, **8**(1), 483–509.

Erdogan, U., Lord, G. J., & Miguez, J. 2025. On a class of constrained Bayesian filters and their numerical implementation in high-dimensional state-space Markov models. *arXiv preprint*, **arXiv:2512.11012**.

Erdogan, U., Lord, G. J., & Miguez, J. 2026. On a class of constrained particle filters for continuous-discrete state space models. *arXiv preprint*, **arXiv:2604.21538**.

Garcia-Fernandez, Angel F, Morelande, Mark R, & Grajal, Jesús. 2012. Truncated unscented Kalman filtering. *IEEE Transactions on Signal Processing*, **60**(7), 3372–3386.

Margellos, Kostas, & Lygeros, John. 2011. Hamilton–Jacobi formulation for reach–avoid differential games. *IEEE Transactions on automatic control*, **56**(8), 1849–1861.

Pieper-Sethmacher, Thorben, Avitabile, Daniele, & van der Meulen, Frank. 2025a. Guided filtering and smoothing for infinite-dimensional diffusions. *arXiv preprint*, **arXiv:2507.06786**.

Pieper-Sethmacher, Thorben, van der Meulen, Frank, & van der Vaart, Aad. 2025b. Simulation of infinite-dimensional diffusion bridges. *arXiv preprint*, **arXiv:2503.13177**.

Rogers, L.C.G., & Williams, D. 2000. *Diffusions, Markov processes, and martingales*. Vol. 2. Cambridge university press.

Wang, Yanbo, Wang, Fasheng, He, Jianjun, & Sun, Fuming. 2020. Iterative Truncated Unscented Particle Filter. *Information (2078-2489)*, **11**(4).

Yang, T., Blom, H., & Mehta, P. G. 2014. The continuous-discrete time feedback particle filter. *Pages 648–653 of: 2014 American Control Conference*. IEEE.