Sequential Monte Carlo methods

Lecture 7 – Auxiliary particle filters

Fredrik Lindsten, Linköping University

2025-02-04

Outline – Lecture 7

Aim: Illustrate the use of "locally optimal" proposals in the auxiliary particle filter (= fully adapted PF)

Outline:

1. Locally optimal proposals
2. When can they be computed?
3. Numerical illustration of fully adapted PF

Fully adapted particle filter

Locally optimal proposals

Recall, auxiliary particle filter

$$
\begin{aligned}
\text{Joint target} \qquad & \propto w_{t-1}^{a_t} p(y_t \mid x_t) p(x_t \mid x_{t-1}^{a_t}) \\
\text{Joint proposal} \qquad & = \nu_{t-1}^{a_t} q(x_t \mid x_{t-1}^{a_t}, y_t)
\end{aligned}
$$

Possible to match term by term?

Locally optimal proposals

Recall, auxiliary particle filter

$$\begin{aligned} &\text{Joint target} && \propto w_{t-1}^{a_t} p(y_t \mid x_t) p(x_t \mid x_{t-1}^{a_t}) \\ &\text{Joint proposal} && = \nu_{t-1}^{a_t} q(x_t \mid x_{t-1}^{a_t}, y_t) \end{aligned}$$

Possible to match term by term? No! — to set proposal = target we first need to **normalize** the target distribution.

Locally optimal proposals

With the choices

Resampling weights: $\nu_{t-1}^i \propto w_{t-1}^i \, p(y_t \mid x_{t-1}^i)$, $i = 1, \ldots, N$

Propagation proposal: $q(x_t \mid x_{t-1}, y_t) = p(x_t \mid x_{t-1}, y_t)$

we are effectively sampling from the joint target for (X_t, A_t).

Locally optimal proposals

With the choices

Resampling weights: $\nu_{t-1}^i \propto w_{t-1}^i \, p(y_t \mid x_{t-1}^i)$, $i = 1, \ldots, N$

Propagation proposal: $q(x_t \mid x_{t-1}, y_t) = p(x_t \mid x_{t-1}, y_t)$

we are effectively sampling from the joint target for (X_t, A_t). $\implies$

$$\widetilde{w}_t^i = \frac{w_{t-1}^{a_t^i}}{\nu_{t-1}^{a_t^i}} \frac{p(y_t \mid x_t^i) p(x_t^i \mid x_{t-1}^{a_t^i})}{q(x_t^i \mid x_{t-1}^{a_t^i}, y_t)} = \text{const.} \quad \implies w_t^i = \frac{1}{N}, \quad i = 1, \ldots, N$$

Locally optimal proposals

With the choices

Resampling weights: $\nu_{t-1}^i \propto w_{t-1}^i \, p(y_t \mid x_{t-1}^i)$, $i = 1, \ldots, N$

Propagation proposal: $q(x_t \mid x_{t-1}, y_t) = p(x_t \mid x_{t-1}, y_t)$

we are effectively sampling from the joint target for (X_t, A_t). $\Longrightarrow$

$$\widetilde{w}_t^i = \frac{w_{t-1}^{a_t^i}}{\nu_{t-1}^{a_t^i}} \frac{p(y_t \mid x_t^i) p(x_t^i \mid x_{t-1}^{a_t^i})}{q(x_t^i \mid x_{t-1}^{a_t^i}, y_t)} = \text{const.} \quad \Longrightarrow w_t^i = \frac{1}{N}, \quad i = 1, \ldots, N$$

Referred to as the fully adapted particle filter (FAPF)

Locally optimal proposals

With the choices

Resampling weights: $\nu_{t-1}^i \quad \propto \quad p(y_t \mid x_{t-1}^i), i = 1, \ldots, N$

Propagation proposal: $q(x_t \mid x_{t-1}, y_t) = p(x_t \mid x_{t-1}, y_t)$

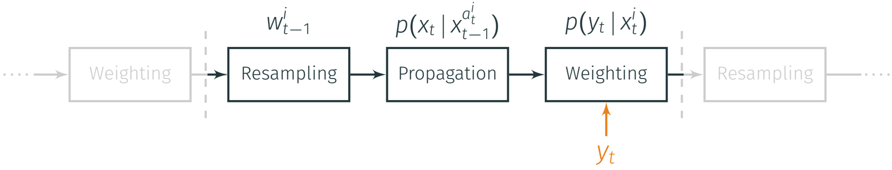
we are effectively sampling from the joint target for (X_t, A_t). $\Longrightarrow$

$$\widetilde{w}_t^i = \frac{w_{t-1}^{a_t^i}}{\nu_{t-1}^{a_t^i}} \frac{p(y_t \mid x_t^i) p(x_t^i \mid x_{t-1}^{a_t^i})}{q(x_t^i \mid x_{t-1}^{a_t^i}, y_t)} = \text{const.} \quad \Longrightarrow w_t^i = \frac{1}{N}, \quad i = 1, \ldots, N$$

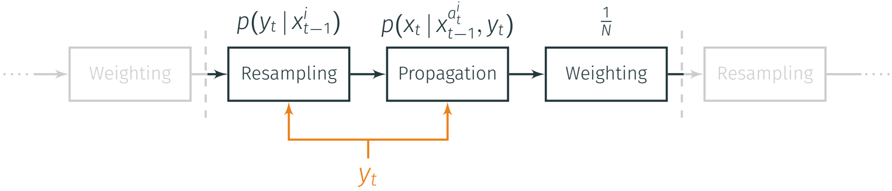
Referred to as the fully adapted particle filter (FAPF)

Locally optimal proposals

Bootstrap particle filter



Fully adapted particle filter



ex) ARCH model

ex) 1st order autoregressive conditional heteroskedasticity (ARCH) model:

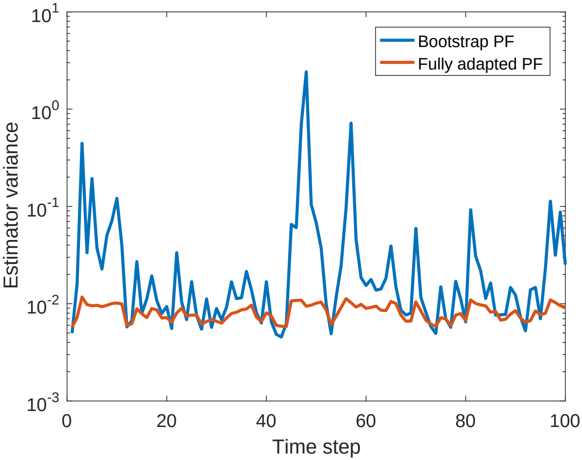
$$X_t = \sqrt{1 + 0.5X_{t-1}^2}V_t, \qquad V_t \sim \mathcal{N}(0,1),$$
$$Y_t = X_t + E_t, \qquad E_t \sim \mathcal{N}(0,r).$$

We simulate a data set and compare the **bootstrap particle filter** with the **fully adapted particle filter**, both using $N = 100$ particles.

Evaluation criteria: Estimator variance for the test function $\varphi(x_t) = x_t$, $t = 1, \ldots, 100$.

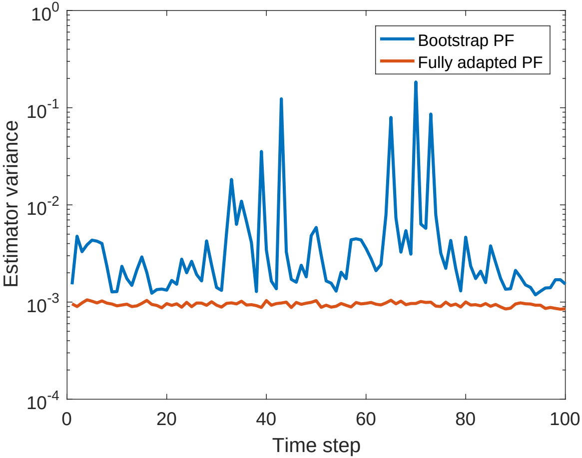
ex) ARCH model

Data set with $r = 1$



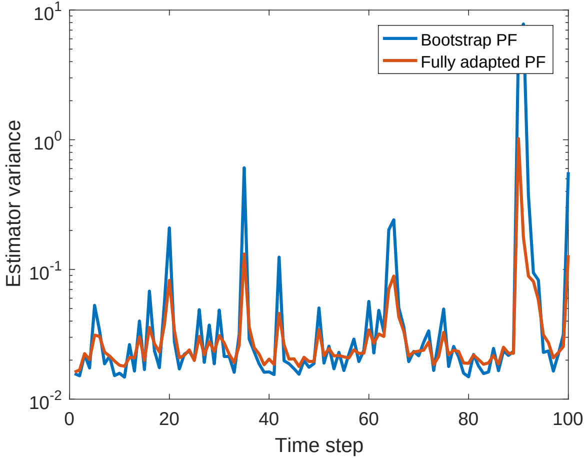
ex) ARCH model

Data set with $r = 0.1$ **(high signal-to-noise ratio)**



ex) ARCH model

Data set with $r = 10$ (low signal-to-noise ratio)

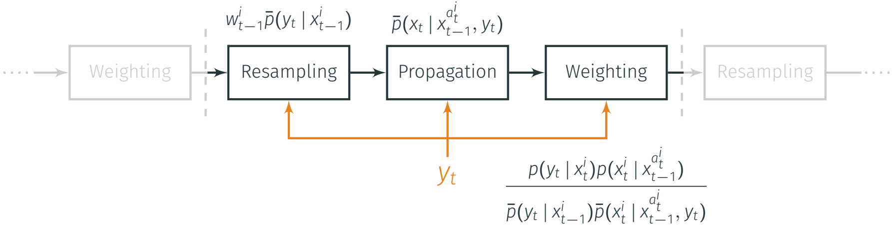


Partially adapted particle filter

Partial adaptation

Non-conjugate models: approximate $\bar{p}(x_t \mid x_{t-1}, y_t) \approx p(x_t \mid x_{t-1}, y_t)$ and $\bar{p}(y_t \mid x_{t-1}) \approx p(y_t \mid x_{t-1})$. E.g., local linearization, variational approximation, ...

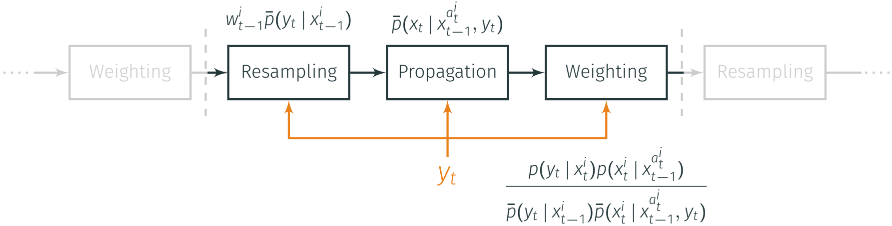
Approximate model used only to **define the proposal!**



Partial adaptation

Non-conjugate models: approximate $\bar{p}(x_t \mid x_{t-1}, y_t) \approx p(x_t \mid x_{t-1}, y_t)$ and $\bar{p}(y_t \mid x_{t-1}) \approx p(y_t \mid x_{t-1})$. E.g., local linearization, variational approximation, ...

Approximate model used only to **define the proposal!**



Care needs to be taken so that the approximations are suitable to use as importance sampling proposals. (Heavier tails than target.)

A few concepts to summarize lecture 7

Locally optimal proposals: Proposals that take all the available information in the current measurement y_t into account.

Fully adapted particle filter: An auxiliary variable that use locally optimal proposals both for the ancestor indices (auxiliary variables) and for the state variable.

Partially adapted particle filter: An auxiliary particle filter that uses some suboptimal proposals (e.g. an approximation of the locally optimal ones) which still take the current measurement y_t into account.