Simple nudging schemes for particle filtering

Ö. Deniz Akyıldız, Joaquín Míguez

Department of Signal Theory and Communications, Universidad Carlos III de Madrid

`omerdeniz.akyildiz@uc3m.es, joaquin.miguez@uc3m.es`

We investigate a new sampling scheme to improve the performance of particle filters in scenarios where either (a) there is a significant mismatch between the assumed model dynamics and the actual system producing the available observations, or (b) the system of interest is high dimensional and the posterior probability tends to concentrate in relatively small regions of the state space.

State-space models and bootstrap particle filters

State-space models (SSMs) consist of a *signal process* $(x_t)_{t\in\mathbb{N}}$ and an *observation process* $(y_t)_{t\in\mathbb{N}}$ with the following conditional independence structure,

$$\begin{aligned} x_0 &\sim \pi_0(x_0) \\ x_t|x_{t-1} &\sim \tau(x_t|x_{t-1}) \\ y_t|x_t &\sim g(y_t|x_t) \end{aligned}$$

where π_0 is the prior distribution, τ is a Markov transition kernel on X, and $g(y_t|x_t)$ denotes the likelihood (we adopt the notation $g_t(x_t) := g(y_t|x_t)$). We are interested in estimating the expectations with respect to conditional distributions $(\pi(x_t|y_{1:t}))_{t\in\mathbb{N}} := (\pi_t(x_t))_{t\in\mathbb{N}}$ sequentially, known as *the filtering problem.*

Suppose we have a set of samples $\{x_{t-1}^{(i)}\}_{i=1}^N$ from $\pi(x_{t-1}|y_{1:t-1})$.

Algorithm 1 (BPF)

For a time step t,

– Sample:

$$\bar{x}_t^{(i)} \sim \tau(x_t|x_{t-1}^{(i)})$$

for all $i = 1, \dots, N$.

– Weight:

$$w_t^{(i)} = g_t(\bar{x}_t^{(i)})/\bar{Z}_t$$

for all $i = 1, \dots, N$ where $\bar{Z}_t = \sum_i g_t(\bar{x}_t^{(i)})$.

– Resample:

$$x_t^{(i)} \sim \sum_{i=1}^N w_t^{(i)} \delta_{\bar{x}_t^{(i)}}(\mathrm{d}x)$$

The samples $\{x_t^{(i)}\}_{i=1}^N$ are used to construct an approximation of $\pi_t \approx \pi_t^N = \frac{1}{N}\sum_i \delta_{x_t^{(i)}}(\mathrm{d}x)$ and used to estimate the integral,

$$\pi_t(f) = \int f(x)\pi_t(\mathrm{d}x) \approx \int f(x)\pi_t^N(\mathrm{d}x) = \frac{1}{N}\sum_i f(x_t^{(i)}) = \pi_t^N(f)$$

BPF performs poorly in high dimensions and under model misspecification.

Nudged particle filter

We propose a modification of BPF, aiming at significantly improving its performance [1]. We use an operation called *nudging* within the particle filter, similar to what has been done in [2]. We define the nudging operator α as,

$$x' = \alpha_t(x) \quad \text{such that} \quad g_t(x') \geq g_t(x) \tag{1}$$

Algorithm 2 (NuPF)

For the time step t,

– Sample:

$$\bar{x}_t^{(i)} \sim \tau(x_t|x_{t-1}^{(i)})$$

for all $i = 1, \dots, N$.

– **Nudge:**

$$\tilde{x}_t^{(i)} = \alpha_t(\bar{x}_t^{(i)})$$

for selected indices $i \in \mathcal{I}$ to be nudged.

– Weight:

$$w_t^{(i)} = g_t(\tilde{x}_t^{(i)})/\tilde{Z}_t$$

for all $i = 1, \dots, N$ where $\tilde{Z}_t = \sum_i g_t(\tilde{x}_t^{(i)})$.

– Resample:

$$x_t^{(i)} \sim \sum_{i=1}^N w_t^{(i)} \delta_{\tilde{x}_t^{(i)}}(\mathrm{d}x)$$

More compactly,

$$\pi_{t-1}^N \underbrace{\rightarrow}_{\text{sampling}} \xi_t^N \underbrace{\rightarrow}_{\text{nudging}} \tilde{\xi}_t^N \underbrace{\rightarrow}_{\text{weighting}} \tilde{\pi}_t^N \underbrace{\rightarrow}_{\text{resampling}} \pi_t^N$$

Nudging schemes

How to choose particles to be nudged: Batch and independent nudging

Batch nudging:

– Choose $\mathcal{I} = \{i_1, \dots, i_M\} \sim [N]$ without replacement.

– $\tilde{x}_t^{(i_k)} = \alpha(\bar{x}_t^{(i_k)})$ for $i_k \in \mathcal{I}$.

Independent nudging:

– $\tilde{x}_t^{(i)} = \alpha(\bar{x}_t^{(i)})$ w. p. M/N.

(For parallel implementations of particle filters)

How to nudge: Gradient nudging

– $\tilde{x}_t^{(i)} = \bar{x}_t^{(i)} + \gamma\nabla_{x_t} g_t(\bar{x}_t^{(i)})$.

There are gradient-free possibilities we do not consider here, see [1].

Analysis

We show that NuPF has the usual Monte Carlo convergence rate, see [1] for more specific results for e.g. gradient steps.

Theorem 1. *Let $y_{1:T}$ be arbitrary but fixed and choose any $0 < t \leq T$. Let f be a bounded test function. Then, under regularity assumptions [1],*

$$\|\pi_t^N(f) - \pi_t(f)\|_p \leq \frac{c_t\|f\|_\infty}{\sqrt{N}}$$

where $c_t < \infty$ is a constant independent of N.

Numerical Experiments

We show results on Lorenz 63 and Lorenz 96 models. For more experiments on object tracking see [1].

Lorenz 63 model

$$\begin{aligned} \mathrm{d}x_1 &= -\mathsf{s}(x_1 - x_2) + \mathrm{d}w_1, \\ \mathrm{d}x_2 &= \mathsf{r}x_1 - x_2 - x_1x_3 + \mathrm{d}w_2, \\ \mathrm{d}x_3 &= x_1x_2 - \mathsf{b}x_3 + \mathrm{d}w_3, \end{aligned}$$

Observations are collected every t_s time steps, with only first dimension is observed. Hence the system is partially observed both time and space,

$$y_n = k_o x_{1,t_s n} + v_n$$

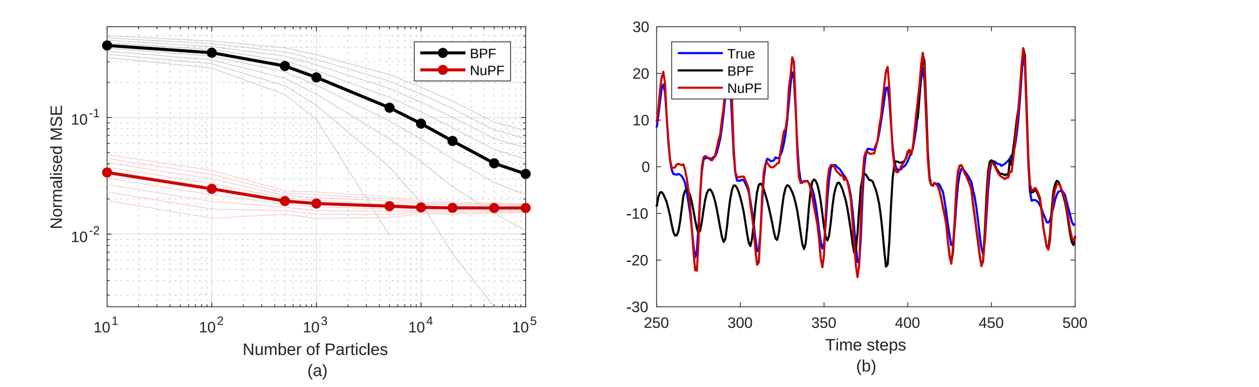

Figure 1. Results with Lorenz 63 misspecified parameter b.

Lorenz 96 model

$$\mathrm{d}x_i = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F + \mathrm{d}w_i$$

Half of the state variables are observed at every t_s time steps,

$$y_{k,n} = x_{(2k-1),t_s n} + v_n$$

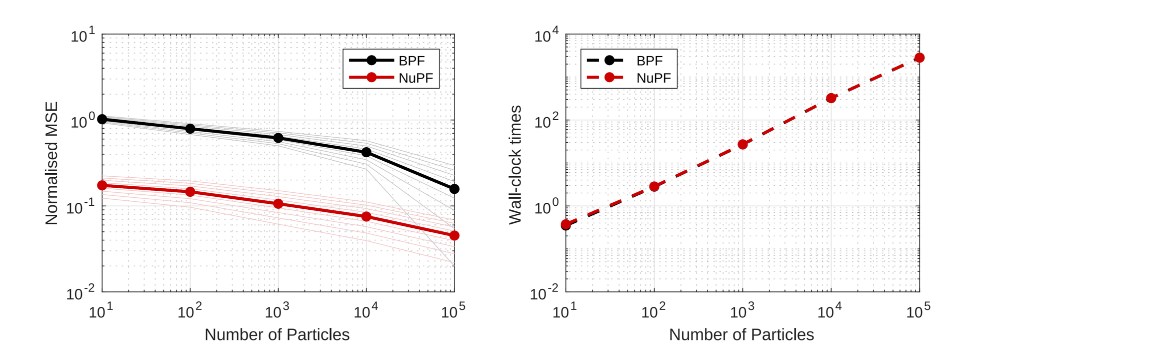

Figure 2. Results with Lorenz 96 model with $d = 40$.

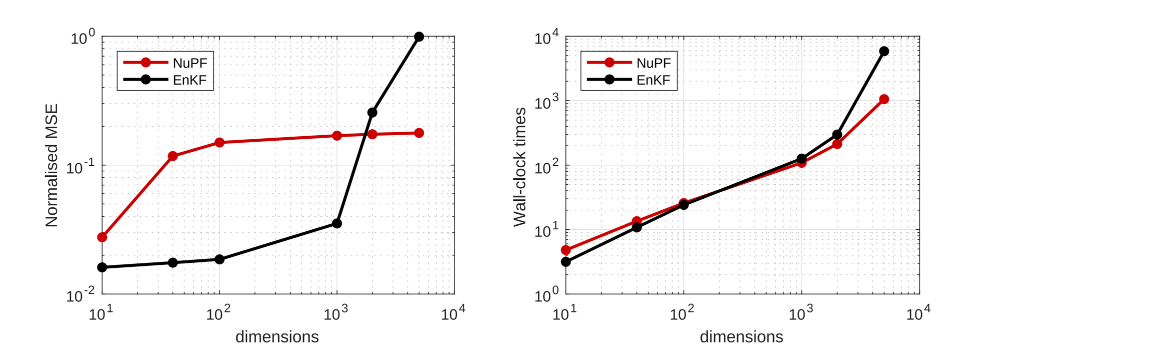

Figure 3. Lorenz 96 model with increasing dimensions with $N = 500$.

References

1. Ö. D. Akyıldız and J. Míguez, "Nudging the Particle Filter", arXiv preprint.

2. M. Ades and P. J. Van Leeuwen. "The equivalent weights particle filter in a high dimensional system." *Quarterly Journal of the Royal Meteorological Society* 141 (687) (2015): 484-503.

Acknowledgements. Ö. D. A. and J. M. acknowledge the support of the Office of Naval Research Global (award no. N62909- 15-1-2011) and *Ministerio de Economía y Competitividad* of Spain (project TEC2015-69868-C2-1-R ADVENTURE).