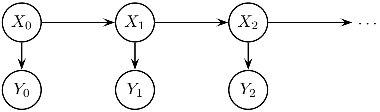
Particle filtering subject to interaction constraints

Nick Whiteley[⋆], Kari Heine[∘], Anthony Lee[+]

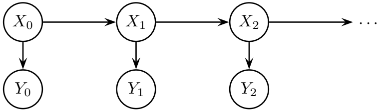
[⋆]University of Bristol, [∘]UCL, [+]University of Warwick,

Hidden Markov Model



$$X_0 \sim \pi_0$$
$$X_n | X_{n-1} \sim f(X_{n-1}, \cdot)$$
$$Y_n | X_n \sim g(X_n, \cdot)$$

Hidden Markov Model

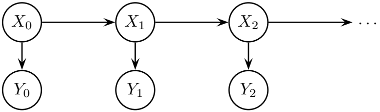


$$X_0 \sim \pi_0$$
$$X_n | X_{n-1} \sim f(X_{n-1}, \cdot)$$
$$Y_n | X_n \sim g(X_n, \cdot)$$

with $\mathbb{E}[\cdot]$ w.r.t. $X_0 \sim \pi_0$, $X_n | X_{n-1} \sim f(X_{n-1}, \cdot)$,

$$\begin{aligned}
\pi_n(A) &:= \mathbb{P}\left(X_n \in A \middle| Y_{0:n-1} = y_{0:n-1} \right) \\
&= \frac{\mathbb{E}\left[\mathbb{I}[X_n \in A] \prod_{k=0}^{n-1} g(X_k, y_k) \right]}{\mathbb{E}\left[\prod_{k=0}^{n-1} g(X_k, y_k) \right]} \\
&= \frac{\int g(x, y_{n-1}) f(x, A) \pi_{n-1}(dx)}{\int g(x, y_{n-1}) \pi_{n-1}(dx)} = \Phi_n(\pi_{n-1})(A)
\end{aligned}$$

Hidden Markov Model



$$X_0 \sim \pi_0$$
$$X_n | X_{n-1} \sim f(X_{n-1}, \cdot)$$
$$Y_n | X_n \sim g(X_n, \cdot)$$

with $\mathbb{E}[\cdot]$ w.r.t. $X_0 \sim \pi_0$, $X_n | X_{n-1} \sim f(X_{n-1}, \cdot)$,

$$\begin{aligned}
\pi_n(A) &:= \mathbb{P}(X_n \in A | Y_{0:n-1} = y_{0:n-1}) \\
&= \frac{\mathbb{E}\left[\mathbb{I}[X_n \in A] \prod_{k=0}^{n-1} g(X_k, y_k)\right]}{\mathbb{E}\left[\prod_{k=0}^{n-1} g(X_k, y_k)\right]} \\
&= \frac{\int g(x, y_{n-1}) f(x, A) \pi_{n-1}(dx)}{\int g(x, y_{n-1}) \pi_{n-1}(dx)} = \Phi_n(\pi_{n-1})(A)
\end{aligned}$$

- $Z_n := \mathbb{E}\left[\prod_{k=0}^{n-1} g(X_k, y_k)\right]$
- write $g_n(x) := g(x, y_n)$

Sequential Importance Sampling

For $n = 0$,

For $i = 1, \ldots, N$,

Set $W_0^i = 1.$

Sample $\zeta_0^i \sim \pi_0,$

For $n \geq 1$,

For $i = 1, \ldots, N$,

Set $W_n^i = W_{n-1}^i g_{n-1}(\zeta_{n-1}^i).$

Sample $\zeta_n^i \sim f(\zeta_{n-1}^i, \cdot).$

Sequential Importance Sampling

For $n = 0$,

  For $i = 1, \ldots, N$,

    Set $\quad W_0^i = 1.$

    Sample $\quad \zeta_0^i \sim \pi_0,$

For $n \geq 1$,

  For $i = 1, \ldots, N$,

    Set $\quad W_n^i = W_{n-1}^i g_{n-1}(\zeta_{n-1}^i).$

    Sample $\quad \zeta_n^i \sim f(\zeta_{n-1}^i, \cdot).$

N.B. the N processes $(\zeta_n^i; n \geq 0)$, $i = 1, \ldots, N$, are independent.

Sequential Importance Sampling

For $n = 0$,

  For $i = 1, \ldots, N$,

    Set $W_0^i = 1$.

    Sample $\zeta_0^i \sim \pi_0$,

For $n \geq 1$,

  For $i = 1, \ldots, N$,

    Set $W_n^i = W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)$.

    Sample $\zeta_n^i \sim f(\zeta_{n-1}^i, \cdot)$.

N.B. the N processes $(\zeta_n^i; n \geq 0)$, $i = 1, \ldots, N$, are independent.

$$\pi_n^N := \frac{\sum_i W_n^i \, \delta_{\zeta_n^i}}{\sum_i W_n^i} \approx \frac{\mathbb{E}\left[\mathbb{I}[X_n \in \cdot] \prod_{k=0}^{n-1} g_k(X_k)\right]}{\mathbb{E}\left[\prod_{k=0}^{n-1} g_k(X_k)\right]} = \pi_n,$$

$$Z_n^N := \frac{1}{N} \sum_i W_n^i \approx \mathbb{E}\left[\prod_{k=0}^{n-1} g_k(X_k)\right] = Z_n.$$

Bootstrap Particle Filter

For $n = 0$,

For $i = 1, \ldots, N$,

Sample $\zeta_0^i \sim \pi_0$,

For $n \geq 1$,

For $i = 1, \ldots, N$,

Sample $$\zeta_n^i \sim \frac{\sum_j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)}{\sum_j g_{n-1}(\zeta_{n-1}^j)}.$$

Bootstrap Particle Filter

For $n = 0$,

For $i = 1, \ldots, N$,

Sample $\zeta_0^i \sim \pi_0$,

For $n \geq 1$,

For $i = 1, \ldots, N$,

Sample $\zeta_n^i \sim \dfrac{\sum_j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)}{\sum_j g_{n-1}(\zeta_{n-1}^j)}$.

$$\pi_n^N := \frac{1}{N} \sum_i \delta_{\zeta_n^i} \qquad Z_n^N := \prod_{p=0}^{n-1} \left(\frac{1}{N} \sum_i g_p(\zeta_p^i) \right)$$

Bootstrap Particle Filter

For $n = 0$,

For $i = 1, \ldots, N$,

Sample $\quad \zeta_0^i \sim \pi_0$,

For $n \geq 1$,

For $i = 1, \ldots, N$,

Sample $\quad \zeta_n^i \;\sim\; \dfrac{\sum_j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)}{\sum_j g_{n-1}(\zeta_{n-1}^j)}$.

$$\pi_n^N := \frac{1}{N} \sum_i \delta_{\zeta_n^i} \qquad Z_n^N := \prod_{p=0}^{n-1} \left(\frac{1}{N} \sum_i g_p(\zeta_p^i) \right)$$

Recall,

$$\pi_n(A) = \frac{\int g_{n-1}(x) f(x, A) \pi_{n-1}(dx)}{\int g_{n-1}(x) \pi_{n-1}(dx)},$$

so notice

$$\frac{\int g_{n-1}(x) f(x, A) \pi_{n-1}^N(dx)}{\int g_{n-1}(x) \pi_{n-1}^N(dx)} = \frac{\sum_j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, A)}{\sum_j g_{n-1}(\zeta_{n-1}^j)}$$

N.B. the N processes $(\zeta_n^i; n \geq 0)$, $i = 1, \ldots, N$, are *not* independent.

Convergence and time-uniform convergence

For both SIS and BPF,

1) $\mathbb{E}[Z_n^N] = Z_n$

2) for bounded $\varphi : \mathbb{X} \to \mathbb{R}$ and $p \geq 1$,

$$\sup_{N \geq 1} \sqrt{N} \mathbb{E}\left[\left| Z_n^N - Z_n \right|^p \right]^{1/p} < +\infty, \quad \sup_{N \geq 1} \sqrt{N} \mathbb{E}\left[\left| \pi_n^N(\varphi) - \pi_n(\varphi) \right|^p \right]^{1/p} < +\infty,$$

so $Z_n^N \to Z_n$ and $\pi_n^N(\varphi) \to \pi_n(\varphi)$ as $N \to \infty$, w.p. 1.

Convergence and time-uniform convergence

For both SIS and BPF,

1) $\mathbb{E}[Z_n^N] = Z_n$

2) for bounded $\varphi : \mathbb{X} \to \mathbb{R}$ and $p \geq 1$,

$$\sup_{N \geq 1} \sqrt{N} \mathbb{E}\left[\left|Z_n^N - Z_n\right|^p\right]^{1/p} < +\infty, \quad \sup_{N \geq 1} \sqrt{N} \mathbb{E}\left[\left|\pi_n^N(\varphi) - \pi_n(\varphi)\right|^p\right]^{1/p} < +\infty,$$

so $Z_n^N \to Z_n$ and $\pi_n^N(\varphi) \to \pi_n(\varphi)$ as $N \to \infty$, w.p. 1.

Highly desirable: for sufficiently regular HMM's,

$$\sup_{N \geq 1} \sup_{n \geq 0} \sqrt{N} \mathbb{E}\left[\left|\pi_n^N(\varphi) - \pi_n(\varphi)\right|^p\right]^{1/p} < +\infty.$$

Convergence and time-uniform convergence

For both SIS and BPF,

1) $\mathbb{E}[Z_n^N] = Z_n$

2) for bounded $\varphi : \mathbb{X} \to \mathbb{R}$ and $p \geq 1$,

$$\sup_{N\geq 1} \sqrt{N}\mathbb{E}\left[\left|Z_n^N - Z_n\right|^p\right]^{1/p} < +\infty, \quad \sup_{N\geq 1} \sqrt{N}\mathbb{E}\left[\left|\pi_n^N(\varphi) - \pi_n(\varphi)\right|^p\right]^{1/p} < +\infty,$$

so $Z_n^N \to Z_n$ and $\pi_n^N(\varphi) \to \pi_n(\varphi)$ as $N \to \infty$, w.p. 1.

Highly desirable: for sufficiently regular HMM's,

$$\sup_{N\geq 1} \sup_{n\geq 0} \sqrt{N}\mathbb{E}\left[\left|\pi_n^N(\varphi) - \pi_n(\varphi)\right|^p\right]^{1/p} < +\infty.$$

(Del Moral and Giuonnet '01) A sufficient condition is: there exists $(\delta, \epsilon) \in [1, \infty)^2$ such that

$$f(x, \cdot) \leq \epsilon f(x', \cdot), \quad \forall x, x', \quad \text{and} \quad \sup_{n\geq 0} \sup_{x,x'} \frac{g_n(x)}{g_n(x')} \leq \delta.$$

Sequential Importance Sampling - quick calculation

For $n = 0$,

For $i = 1, \ldots, N$,

Set $W_0^i = 1.$

Sample $\zeta_0^i \sim \pi_0,$

For $n \geq 1$,

For $i = 1, \ldots, N$,

Set $W_n^i = W_{n-1}^i g_{n-1}(\zeta_{n-1}^i).$

Sample $\zeta_n^i \sim f(\zeta_{n-1}^i, \cdot).$

$$\sqrt{N}\left[\pi_n^N(\varphi) - \pi_n(\varphi)\right] = \sqrt{N}\left[\frac{\sum_i W_n^i\, \varphi(\zeta_n^i)}{\sum_i W_n^i} - \pi_n(\varphi)\right] = \frac{\frac{1}{\sqrt{N}}\sum_i W_n^i\ \left[\varphi(\zeta_n^i) - \pi_n(\varphi)\right]}{\frac{1}{N}\sum_i W_n^i}$$

Sequential Importance Sampling - quick calculation

For $n = 0$,
For $i = 1, \ldots, N$,
Set $W_0^i = 1$.
Sample $\zeta_0^i \sim \pi_0$,
For $n \geq 1$,
For $i = 1, \ldots, N$,
Set $W_n^i = W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)$.
Sample $\zeta_n^i \sim f(\zeta_{n-1}^i, \cdot)$.

$$\sqrt{N}\left[\pi_n^N(\varphi) - \pi_n(\varphi)\right] = \sqrt{N}\left[\frac{\sum_i W_n^i\, \varphi(\zeta_n^i)}{\sum_i W_n^i} - \pi_n(\varphi)\right] = \frac{\frac{1}{\sqrt{N}} \sum_i W_n^i \left[\varphi(\zeta_n^i) - \pi_n(\varphi)\right]}{\frac{1}{N} \sum_i W_n^i}$$

- Numer. obeys CLT, denom. converges a.s. to $\mathbb{E}[\prod_{k=0}^{n-1} g_k(X_k)]$, so

$$\sqrt{N}\left[\pi_n^N(\varphi) - \pi_n(\varphi)\right] \xrightarrow[N\to\infty]{d} \mathcal{N}(0, \sigma_n^2), \qquad \sigma_n^2 = \frac{\mathbb{E}\left[(\varphi(X_n) - \pi_n(\varphi))^2 \prod_{k=0}^{n-1} g_k(X_k)^2\right]}{\mathbb{E}\left[\prod_{k=0}^{n-1} g_k(X_k)\right]^2}.$$

Sequential Importance Sampling - quick calculation

For $n = 0$,
  For $i = 1, \ldots, N$,
    Set $W_0^i = 1$.
    Sample $\zeta_0^i \sim \pi_0$,
For $n \geq 1$,
  For $i = 1, \ldots, N$,
    Set $W_n^i = W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)$.
    Sample $\zeta_n^i \sim f(\zeta_{n-1}^i, \cdot)$.

$$\sqrt{N}\left[\pi_n^N(\varphi) - \pi_n(\varphi)\right] = \sqrt{N}\left[\frac{\sum_i W_n^i\, \varphi(\zeta_n^i)}{\sum_i W_n^i} - \pi_n(\varphi)\right] = \frac{\frac{1}{\sqrt{N}} \sum_i W_n^i\, \left[\varphi(\zeta_n^i) - \pi_n(\varphi)\right]}{\frac{1}{N}\sum_i W_n^i}$$

- Numer. obeys CLT, denom. converges a.s. to $\mathbb{E}[\prod_{k=0}^{n-1} g_k(X_k)]$, so

$$\sqrt{N}\left[\pi_n^N(\varphi) - \pi_n(\varphi)\right] \xrightarrow[N\to\infty]{d} \mathcal{N}(0, \sigma_n^2), \qquad \sigma_n^2 = \frac{\mathbb{E}\left[(\varphi(X_n) - \pi_n(\varphi))^2 \prod_{k=0}^{n-1} g_k(X_k)^2\right]}{\mathbb{E}\left[\prod_{k=0}^{n-1} g_k(X_k)\right]^2}.$$

- If $f(x, \cdot) = \pi_0(\cdot)$, and $g_k = g$, then

$$\sigma_n^2 = \mathrm{Var}_{\pi_0}(\varphi)\left(\frac{\pi_0(g^2)}{\pi_0(g)^2}\right)^n.$$

Overview

- Part I: analysis of a "distributed" particle filter
  - Bolić et al., *Resampling Algorithms and Architectures for Distributed Particle filters*, IEEE Trans. Sig. Proc. 2005.
  - *Heine and W., ArXiv imminent*
- Part II: how much interaction is enough?
  - *On the role of interaction in sequential Monte Carlo algorithms*. W, Lee and Heine. Bernoulli. To appear.

Part I: analysis of a distributed particle filter

Local exchange particle filter (Bolić et al., 2005)

- $N = Mm$ particles in total, $m \equiv$ no. of "machines"
- $\theta \in \{1, \dots, M-1\}$ is a parameter
- write $y \bmod^\star x = y - \lfloor (y-1)/x \rfloor x$.

For $n = 0$,
  For $i = 1, \dots, N$,
    Set $W_0^i = 1$ and sample $\zeta_0^i \sim \pi_0$
    Set $L^i = (i + \theta) \bmod^\star Mm$,
  For $k = 1, \dots, m$
    Set $G_k = \{(k-1)M + 1, \dots, (k-1)M + M\}$
For $n \geq 1$,
  For $k = 1, \dots, m$,
    For $i \in G_k$,
      Set $W_n^i = (Mm)^{-1} \sum_{j \in G_k} W_{n-1}^{L^j} g_{n-1}(\zeta_{n-1}^{L^j})$.
      Sample $\zeta_n^i \sim (W_n^i)^{-1} \sum_{j \in G_k} W_{n-1}^{L^j} g_{n-1}(\zeta_{n-1}^{L^j}) f(\zeta_{n-1}^{L^j}, \cdot)$

Local exchange particle filter (Bolić et al., 2005)

- $N = Mm$ particles in total, $m \equiv$ no. of "machines"
- $\theta \in \{1, \dots, M-1\}$ is a parameter
- write $y \text{mod}^\star x = y - \lfloor (y-1)/x \rfloor x$.

For $n = 0$,
  For $i = 1, \dots, N$,
    Set $W_0^i = 1$ and sample $\zeta_0^i \sim \pi_0$
    Set $L^i = (i + \theta)\text{mod}^\star Mm$,
  For $k = 1, \dots, m$
    Set $G_k = \{(k-1)M + 1, \dots, (k-1)M + M\}$
For $n \geq 1$,
  For $k = 1, \dots, m$,
    For $i \in G_k$,
      Set $W_n^i = (Mm)^{-1} \sum_{j \in G_k} W_{n-1}^{L^j} g_{n-1}(\zeta_{n-1}^{L^j})$.
      Sample $\zeta_n^i \sim (W_n^i)^{-1} \sum_{j \in G_k} W_{n-1}^{L^j} g_{n-1}(\zeta_{n-1}^{L^j}) f(\zeta_{n-1}^{L^j}, \cdot)$

$$\pi_n^N := \frac{\sum_i W_n^i \, \delta_{\zeta_n^i}}{\sum_i W_n^i} \qquad Z_n^N := \frac{1}{N} \sum_i W_n^i$$

with $\theta = 0$, LEPF $\equiv$ independent bootstrap particle filters

For $n = 0$,

For $i = 1, \ldots, N$,

Set $W_0^i = 1$ and sample $\zeta_0^i \sim \pi_0$

For $k = 1, \ldots, m$

Set $G_k = \{(k-1)M + 1, \ldots, (k-1)M + M\}$

For $n \geq 1$,

For $k = 1, \ldots, m$,

For $i \in G_k$,

Set $W_n^i = (Mm)^{-1} \sum_{j \in G_k} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)$.

Sample $\zeta_n^i \sim (W_n^i)^{-1} \sum_{j \in G_k} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)$

with $\theta = 0$, LEPF $\equiv$ independent bootstrap particle filters

---

For $n = 0$,
  For $i = 1, \ldots, N$,
    Set $W_0^i = 1$ and sample $\zeta_0^i \sim \pi_0$
  For $k = 1, \ldots, m$
    Set $G_k = \{(k-1)M + 1, \ldots, (k-1)M + M\}$
For $n \geq 1$,
  For $k = 1, \ldots, m$,
    For $i \in G_k$,
      Set $W_n^i = (Mm)^{-1} \sum_{j \in G_k} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)$.
      Sample $\zeta_n^i \sim (W_n^i)^{-1} \sum_{j \in G_k} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)$

- N.B. $W_n^i = W_n^j$ for all $i, j \in G_k$, and

$$\mathbb{P}(\zeta_n^i \in \cdot | \zeta_0, \ldots, \zeta_{n-1}) = \frac{\sum_{j \in G_k} g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)}{\sum_{j \in G_k} g_{n-1}(\zeta_{n-1}^j)}.$$

A general algorithm - αSMC

- α is a $N \times N$ row-stochastic matrix

For $n = 0$,
  For $i = 1, \ldots, N$,
    Set $W_0^i = 1$ and sample $\zeta_0^i \sim \pi_0$
For $n \geq 1$,
  For $i = 1, \ldots, N$,
    Set $W_n^i = \sum_j \alpha^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)$
    Sample $\zeta_n^i \sim (W_n^i)^{-1} \sum_j \alpha^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)$

A general algorithm - αSMC

- α is a $N \times N$ row-stochastic matrix

For $n = 0$,
  For $i = 1, \dots, N$,
    Set $W_0^i = 1$ and sample $\zeta_0^i \sim \pi_0$
For $n \geq 1$,
  For $i = 1, \dots, N$,
    Set $W_n^i = \sum_j \alpha^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)$
    Sample $\zeta_n^i \sim (W_n^i)^{-1} \sum_j \alpha^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)$

$$\pi_n^N := \frac{\sum_i W_n^i \, \delta_{\zeta_n^i}}{\sum_i W_n^i} \qquad Z_n^N := \frac{1}{N} \sum_i W_n^i$$

LEPF and IBPF as αSMC

For $n = 0$,

  For $i = 1, \ldots, N$,

    Set $W_0^i = 1$ and sample $\zeta_0^i \sim \pi_0$

For $n \geq 1$,

  For $i = 1, \ldots, N$,

    Set $W_n^i = \sum_j \alpha^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)$

    Sample $\zeta_n^i \sim (W_n^i)^{-1} \sum_j \alpha^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)$

$$\begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 & 0 \\ 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 \\ 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 \\ 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \end{pmatrix}$$

$m = M = 3$. left: α for LEPF with $\theta = 1$, right: α for IBPF.

LEPF and IBPF as αSMC

For $n = 0$,

  For $i = 1, \ldots, N$,

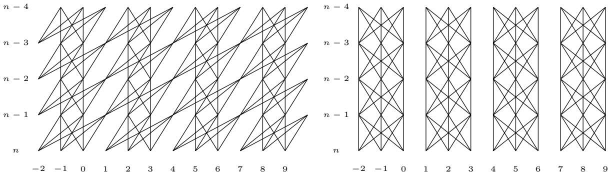
    Set $\quad W_0^i = 1$ and sample $\quad \zeta_0^i \sim \pi_0$

For $n \geq 1$,

  For $i = 1, \ldots, N$,

    Set $\quad W_n^i = \sum_j \alpha^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)$

    Sample $\quad \zeta_n^i \sim (W_n^i)^{-1} \sum_j \alpha^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)$



$M = 3$. left: α for LEPF with $\theta = 1$, right: α for IBPF.

Convergence and CLT

Thm. The following hold for the LEPF and IBPF

1) For any $M, m \geq 1$, $\mathbb{E}[Z_n^{Mm}] = Z_n$

2) for bounded $\varphi : \mathbb{X} \to \mathbb{R}$ and $p \geq 1$,

$$\sup_{M,m\geq 1} \sqrt{Mm}\mathbb{E}\left[\left|Z_n^{Mm} - Z_n\right|^p\right]^{1/p} < +\infty$$

$$\sup_{M,m\geq 1} \sqrt{Mm}\mathbb{E}\left[\left|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)\right|^p\right]^{1/p} < +\infty$$

so $Z_n^{Mm} \to Z_n$ and $\pi_n^{Mm}(\varphi) \to \pi_n(\varphi)$, w.p. 1. both:

- with m fixed and $M \to \infty$, and
- with M fixed and $m \to \infty$

3) for bounded $\varphi : \mathbb{X} \to \mathbb{R}$ and any $M \geq 2$,

$$\sqrt{Mm}\left(\pi_n^{Mm}(\varphi) - \pi_n(\varphi)\right) \xrightarrow[m\to\infty]{d} \mathcal{N}(0, \sigma_n^2),$$

with σ_n^2 depending on φ, M (and θ in the case of the LEPF)

Time-uniform convergence

Thm. The following hold for the IBPF and LEPF

1) If there exists $(\delta, \epsilon) \in [1, \infty)^2$ such that

$$f(x, \cdot) \leq \epsilon f(x', \cdot), \quad \forall x, x', \quad \text{and} \quad \sup_{n \geq 0} \sup_{x, x'} \frac{g_n(x)}{g_n(x')} \leq \delta, \tag{1}$$

then for any $m \geq 1$, $p \geq 1$ and bounded $\varphi : \mathbb{X} \to \mathbb{R}$,

$$\sup_{M \geq 1} \sup_{n \geq 0} \sqrt{M} \mathbb{E} \left[\left| \pi_n^{Mm}(\varphi) - \pi_n(\varphi) \right|^p \right]^{1/p} < +\infty$$

Time-uniform convergence

Thm. The following hold for the IBPF and LEPF

1) If there exists $(\delta, \epsilon) \in [1, \infty)^2$ such that

$$f(x, \cdot) \leq \epsilon f(x', \cdot), \quad \forall x, x', \quad \text{and} \quad \sup_{n\geq 0} \sup_{x,x'} \frac{g_n(x)}{g_n(x')} \leq \delta, \tag{1}$$

then for any $m \geq 1$, $p \geq 1$ and bounded $\varphi : \mathbb{X} \to \mathbb{R}$,

$$\sup_{M\geq 1} \sup_{n\geq 0} \sqrt{M}\mathbb{E}\left[\left|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)\right|^p\right]^{1/p} < +\infty$$

2) but, there exist models satisying (1) such that for any φ such that $\mathrm{Var}_{\pi_0}(\varphi) > 0$, and any $M \geq 1$ and $p \geq 1$,

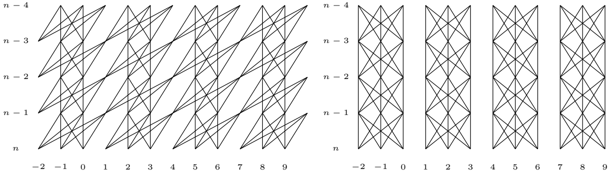
$$\lim_{m\to\infty} \sup_{n\geq 0} \sqrt{m}\mathbb{E}\left[\left|\pi_n^{Mm}(\varphi) - \pi_n(\varphi)\right|^p\right]^{1/p} = +\infty.$$

to establish 2), it turns out to be sufficient to construct a model such that $\lim_{n\to\infty} \sigma_n^2 = +\infty$...

Asymptotic variance in a simplified case

- Introduce the "limiting" α matrix for the LEPF (for IBPF set $\theta = 0$):

$$\alpha_\infty^{ij} := M^{-1}\mathbb{I}[\lfloor (i-1)/M \rfloor = \lfloor (j-\theta-1)/M \rfloor], \quad i,j \in \mathbb{Z}$$



- Let $(I_k, J_k)_{0 \le k \le n}$ be the bi-variate, backward Markov chain:

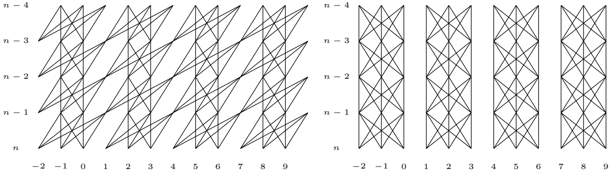
$$(I_n, J_n) \sim \delta_0 \otimes \delta_0, \quad \mathbb{P}(I_k = i_k, J_k = j_k | I_{k+1} = i_{k+1}, J_{k+1} = j_{k+1}) = \alpha_\infty^{i_{k+1} i_k} \alpha_\infty^{j_{k+1} j_k},$$

- Define the collision count: $Z_n = \sum_{k=0}^{n-1} \mathbb{I}[I_k = J_k]$

Asymptotic variance in a simplified case

- Introduce the "limiting" α matrix for the LEPF (for IBPF set $\theta = 0$):

$$\alpha_\infty^{ij} := M^{-1}\mathbb{I}[\lfloor (i-1)/M \rfloor = \lfloor (j-\theta-1)/M \rfloor], \quad i,j \in \mathbb{Z}$$



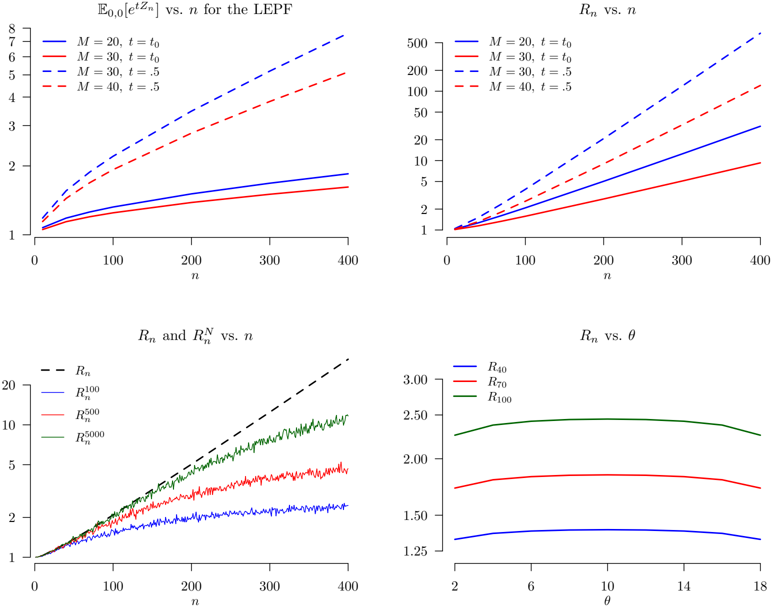
- Let $(I_k, J_k)_{0 \le k \le n}$ be the bi-variate, backward Markov chain:

$$(I_n, J_n) \sim \delta_0 \otimes \delta_0, \quad \mathbb{P}(I_k = i_k, J_k = j_k | I_{k+1} = i_{k+1}, J_{k+1} = j_{k+1}) = \alpha_\infty^{i_{k+1} i_k} \alpha_\infty^{j_{k+1} j_k},$$

- Define the collision count: $Z_n = \sum_{k=0}^{n-1} \mathbb{I}[I_k = J_k]$
- Assume $f(x, \cdot) = \pi_0(\cdot)$ and $g_n = g$, then

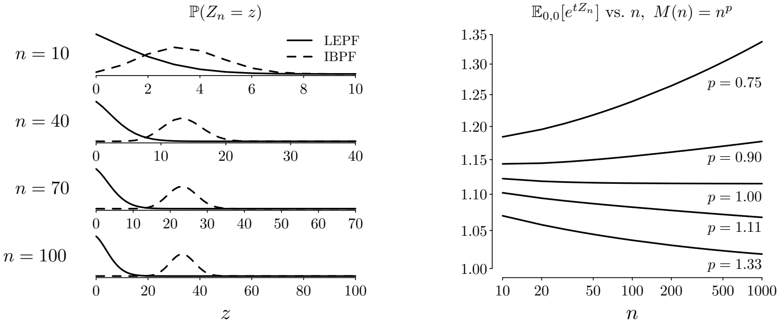
$$\sigma_n^2 = \mathrm{Var}_{\pi_0}(\varphi)\mathbb{E}_{0,0}\left[e^{tZ_n}\right], \quad t := \log \frac{\pi_0(g^2)}{\pi_0(g)^2}.$$

Numerical comparisons

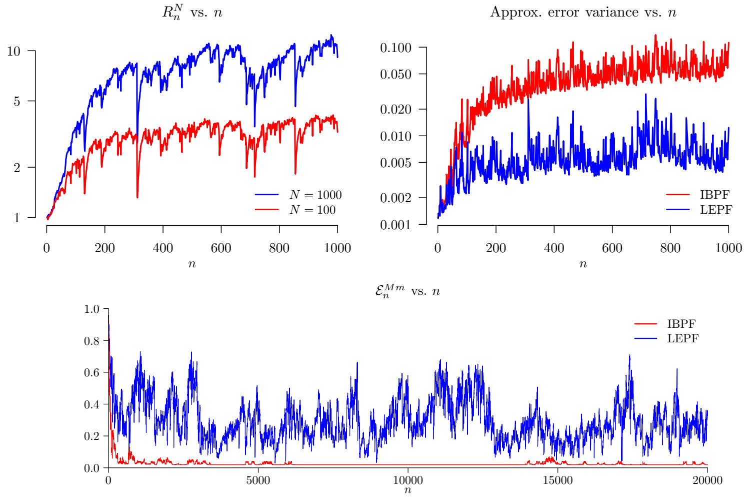


Numerical comparisons

- $Z_n = \sum_{k=0}^{n-1} \mathbb{I}[I_k = J_k]$
- For the IBPF, $Z_n \sim \text{Binomial}(n, M^{-1})$
- $\sigma_n^2 = \text{Var}_{\pi_0}(\varphi)\mathbb{E}_{0,0}\left[e^{tZ_n}\right], \quad t := \log \frac{\pi_0(g^2)}{\pi_0(g)^2}.$



Simulation example



$$\mathcal{E}_n^N = \frac{\left(N^{-1}\sum_{i=1}^N W_n^i\right)^2}{N^{-1}\sum_{i=1}^N (W_n^i)^2} \in [0,1], \qquad \pi_n^N = \frac{\sum_{i=1}^N W_n^i \delta_{\zeta_n^i}}{\sum_{i=1}^N W_n^i}.$$

Part II: how much interaction is enough?

αSMC - a general algorithm

Let $\mathbb{A}_N$ be a set of Markov transition matrices, each of size $N \times N$

For $n = 0$,
 For $i = 1, \ldots, N$,
  Set $W_0^i = 1$
  Sample $\zeta_0^i \sim \pi_0$
For $n \geq 1$,
 Select α_{n-1} from $\mathbb{A}_N$ according to some given functional of $\{\zeta_0, \ldots, \zeta_{n-1}\}$
 For $i = 1, \ldots, N$,
  Set $W_n^i = \sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)$
  Sample $\zeta_n^i \sim \dfrac{\sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)}{W_n^i}$

$$\pi_n^N := \frac{\sum_i W_n^i \, \delta_{\zeta_n^i}}{\sum_i W_n^i} \qquad Z_n^N := \frac{1}{N} \sum_i W_n^i$$

αSMC - a general algorithm

Let $\mathbb{A}_N$ be a set of Markov transition matrices, each of size $N \times N$

For $n = 0$,
  For $i = 1, \ldots, N$,
    Set $W_0^i = 1$
    Sample $\zeta_0^i \sim \pi_0$
For $n \geq 1$,
  Select α_{n-1} from $\mathbb{A}_N$ according to some given functional of $\{\zeta_0, \ldots, \zeta_{n-1}\}$
  For $i = 1, \ldots, N$,
    Set $W_n^i = \sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)$
    Sample $\zeta_n^i \sim \dfrac{\sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)}{W_n^i}$

$$\pi_n^N := \frac{\sum_i W_n^i \, \delta_{\zeta_n^i}}{\sum_i W_n^i} \qquad Z_n^N := \frac{1}{N} \sum_i W_n^i$$

- If $\mathbb{A}_N = \{Id\}$, αSMC $\equiv$ sequential importance sampling
- If $\mathbb{A}_N = \{\mathbf{1}_{1/N}\}$, αSMC $\equiv$ bootstrap particle filter

($\mathbf{1}_{1/N}$ is the matrix with $1/N$ as every entry)

Adaptive Resampling PF (Liu and Chen '95) as αSMC

- Adaptive Resampling PF is an instance of αSMC with $\tau \in (0,1]$ a threshold, $\mathbb{A}_N = \left\{ Id, \mathbf{1}_{1/N} \right\}$, and

$$\alpha_{n-1} := \begin{cases} \mathbf{1}_{1/N}, & \text{if} \quad \dfrac{\left(N^{-1}\sum_i W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)\right)^2}{N^{-1}\sum_i \left(W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)\right)^2} < \tau. \\ Id, & \text{otherwise.} \end{cases}$$

Adaptive Resampling PF (Liu and Chen '95) as αSMC

- Adaptive Resampling PF is an instance of αSMC with $\tau \in (0,1]$ a threshold, $\mathbb{A}_N = \left\{ Id, \mathbf{1}_{1/N} \right\}$, and

$$\alpha_{n-1} := \begin{cases} \mathbf{1}_{1/N}, & \text{if} \quad \frac{\left(N^{-1}\sum_i W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)\right)^2}{N^{-1}\sum_i \left(W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)\right)^2} < \tau. \\ Id, & \text{otherwise.} \end{cases}$$

- For general αSMC, introduce the key quantity:

$$\mathcal{E}_n^N := \frac{(N^{-1}\sum_i W_n^i)^2}{N^{-1}\sum_i (W_n^i)^2} = \frac{\left(N^{-1}\sum_i \sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)\right)^2}{N^{-1}\sum_i \left(\sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)\right)^2} \in [0,1]$$

Adaptive Resampling PF (Liu and Chen '95) as αSMC

- Adaptive Resampling PF is an instance of αSMC with $\tau \in (0,1]$ a threshold, $\mathbb{A}_N = \left\{ Id, \mathbf{1}_{1/N} \right\}$, and

$$\alpha_{n-1} := \begin{cases} \mathbf{1}_{1/N}, & \text{if} \quad \dfrac{\left(N^{-1}\sum_i W^i_{n-1} g_{n-1}(\zeta^i_{n-1})\right)^2}{N^{-1}\sum_i \left(W^i_{n-1} g_{n-1}(\zeta^i_{n-1})\right)^2} < \tau. \\ Id, & \text{otherwise.} \end{cases}$$

- For general αSMC, introduce the key quantity:

$$\mathcal{E}^N_n := \frac{\left(N^{-1}\sum_i W^i_n\right)^2}{N^{-1}\sum_i (W^i_n)^2} = \frac{\left(N^{-1}\sum_i \sum_j \alpha^{ij}_{n-1} W^j_{n-1} g_{n-1}(\zeta^j_{n-1})\right)^2}{N^{-1}\sum_i \left(\sum_j \alpha^{ij}_{n-1} W^j_{n-1} g_{n-1}(\zeta^j_{n-1})\right)^2} \in [0,1]$$

- Notice

$$\alpha_{n-1} = \mathbf{1}_{1/N} \quad \Rightarrow \quad \mathcal{E}^N_n = 1,$$

$$\alpha_{n-1} = Id \quad \Rightarrow \quad \mathcal{E}^N_n = \frac{\left(N^{-1}\sum_i W^i_{n-1} g_{n-1}(\zeta^i_{n-1})\right)^2}{N^{-1}\sum_i \left(W^i_{n-1} g_{n-1}(\zeta^i_{n-1})\right)^2}.$$

Adaptive Resampling PF (Liu and Chen '95) as αSMC

- Adaptive Resampling PF is an instance of αSMC with $\tau \in (0,1]$ a threshold, $\mathbb{A}_N = \left\{ Id, \mathbf{1}_{1/N} \right\}$, and

$$\alpha_{n-1} := \begin{cases} \mathbf{1}_{1/N}, & \text{if } \quad \dfrac{\left(N^{-1}\sum_i W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)\right)^2}{N^{-1}\sum_i \left(W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)\right)^2} < \tau. \\ Id, & \text{otherwise.} \end{cases}$$

- For general αSMC, introduce the key quantity:

$$\mathcal{E}_n^N := \frac{(N^{-1}\sum_i W_n^i)^2}{N^{-1}\sum_i (W_n^i)^2} = \frac{\left(N^{-1}\sum_i \sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)\right)^2}{N^{-1}\sum_i \left(\sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j)\right)^2} \in [0,1]$$

- Notice

$$\alpha_{n-1} = \mathbf{1}_{1/N} \quad \Rightarrow \quad \mathcal{E}_n^N = 1,$$

$$\alpha_{n-1} = Id \quad \Rightarrow \quad \mathcal{E}_n^N = \frac{\left(N^{-1}\sum_i W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)\right)^2}{N^{-1}\sum_i \left(W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)\right)^2}.$$

- So ARPF can be viewed as enforcing:

$$\inf_n \mathcal{E}_n^N \geq \tau$$

Convergence of αSMC

(A) - the entries of the $\mathbb{A}_N$-valued random matrix α_{n-1} are measurable w.r.t. $\mathcal{F}_{n-1} = \sigma(\zeta_0, ..., \zeta_{n-1})$

(B) - every member of $\mathbb{A}_N$ is doubly stochastic

Convergence of αSMC

(**A**) - the entries of the $\mathbb{A}_N$-valued random matrix α_{n-1} are measurable w.r.t. $\mathcal{F}_{n-1} = \sigma(\zeta_0, ..., \zeta_{n-1})$

(**B**) - every member of $\mathbb{A}_N$ is doubly stochastic

---

Thm. Assume (**A**) and (**B**). Then for any $n \geq 0$,

1) $\mathbb{E}[Z_n^N] = Z_n$

2) for bounded $\varphi : \mathbb{X} \to \mathbb{R}$ and $p \geq 1$,

$$\sup_{N\geq 1} \sqrt{N}\mathbb{E}\left[\left|Z_n^N - Z_n\right|^p\right]^{1/p} < +\infty, \quad \sup_{N\geq 1} \sqrt{N}\mathbb{E}\left[\left|\pi_n^N(\varphi) - \pi_n(\varphi)\right|^p\right]^{1/p} < +\infty,$$

so $Z_n^N \to Z_n$ and $\pi_n^N(\varphi) \to \pi_n(\varphi)$ as $N \to \infty$, w.p. 1.

Uniform-convergence of αSMC

(**C**) There exists $(\delta, \epsilon) \in [1, \infty)^2$ such that

$$f(x, \cdot) \leq \epsilon f(x', \cdot), \quad \forall x, x', \qquad \text{and} \qquad \sup_{n \geq 0} \sup_{x, x'} \frac{g_n(x)}{g_n(x')} \leq \delta.$$

Uniform-convergence of αSMC

(**C**) There exists $(\delta, \epsilon) \in [1, \infty)^2$ such that

$$f(x, \cdot) \leq \epsilon f(x', \cdot), \quad \forall x, x', \quad \text{and} \quad \sup_{n \geq 0} \sup_{x, x'} \frac{g_n(x)}{g_n(x')} \leq \delta.$$

Thm. Assume (**A**), (**B**) and (**C**). Then there exist finite constants, c_1 and for each $p \geq 1$, $c_2(p)$, such that for any $\tau \in (0, 1]$, $N \geq 1$, and bounded $\varphi : \mathbb{X} \to \mathbb{R}$

$$\inf_{n \geq 0} \mathcal{E}_n^N \geq \tau \quad \Rightarrow \quad \begin{cases} \sup_{n \geq 1} \mathbb{E}\left[\left(\frac{Z_n^N}{Z_n}\right)^2\right]^{1/n} \leq 1 + \frac{c_1}{N\tau}, \\ \text{and} \\ \sup_{n \geq 0} \mathbb{E}\left[\left|\pi_n^N(\varphi) - \pi_n(\varphi)\right|^p\right]^{1/p} \leq \|\varphi\| \frac{c_2(p)}{\sqrt{N\tau}}. \end{cases}$$

Uniform-convergence of αSMC

(C) There exists $(\delta, \epsilon) \in [1, \infty)^2$ such that

$$f(x, \cdot) \leq \epsilon f(x', \cdot), \quad \forall x, x', \quad \text{and} \quad \sup_{n \geq 0} \sup_{x, x'} \frac{g_n(x)}{g_n(x')} \leq \delta.$$

Thm. Assume **(A)**, **(B)** and **(C)**. Then there exist finite constants, c_1 and for each $p \geq 1$, $c_2(p)$, such that for any $\tau \in (0, 1]$, $N \geq 1$, and bounded $\varphi : \mathbb{X} \to \mathbb{R}$

$$\inf_{n \geq 0} \mathcal{E}_n^N \geq \tau \quad \Rightarrow \quad \begin{cases} \sup_{n \geq 1} \mathbb{E}\left[\left(\frac{Z_n^N}{Z_n}\right)^2\right]^{1/n} \leq 1 + \frac{c_1}{N\tau}, \\ \text{and} \\ \sup_{n \geq 0} \mathbb{E}\left[\left|\pi_n^N(\varphi) - \pi_n(\varphi)\right|^p\right]^{1/p} \leq \|\varphi\| \frac{c_2(p)}{\sqrt{N\tau}}. \end{cases}$$

Cor.

$$\inf_{n \geq 0} \mathcal{E}_n^N \geq \tau \text{ and } N\tau \geq nc_1 \quad \Rightarrow \quad \mathbb{E}\left[\left(\frac{Z_n^N}{Z_n} - 1\right)^2\right] \leq \frac{2nc_1}{N\tau}.$$

Design of αSMC

	SIS	PF	Adaptive Resampling PF	new instances of αSMC
$\mathbb{A}_N$	$\{Id\}$	$\{\mathbf{1}_{1/N}\}$	$\{Id, \mathbf{1}_{1/N}\}$	a matter of design
α_n chosen from $\mathbb{A}_N$ according to some function of $\{\zeta_0, \ldots, \zeta_n\}$	✓	✓	✓	easy to verify
$\mathbb{A}_N$ contains only doubly stochastic matrices	✓	✓	✓	easy to verify
$\inf_{n \geq 0} \mathcal{E}_n^N > 0$ enforced	×	✓	✓	a matter of design

Example of adaptive interaction

- assume $N = 2^m$, $m \in \mathbb{N}$ and let

$$\mathbb{A}_N = \left\{ Id, a(1), ..., a(m-1), \mathbf{1}_{1/N} \right\},$$

with

$$a(k) = \begin{bmatrix} \mathbf{1}_{1/2^k} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{1/2^k} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1}_{1/2^k} \end{bmatrix}, \quad k = 0, ..., m,$$

where $\mathbf{1}_{1/d}$ is the $d \times d$ matrix with $1/d$ as every entry.

Example of adaptive interaction

- assume $N = 2^m$, $m \in \mathbb{N}$ and let

$$\mathbb{A}_N = \left\{ Id, a(1), ..., a(m-1), \mathbf{1}_{1/N} \right\},$$

with

$$a(k) = \begin{bmatrix} \mathbf{1}_{1/2^k} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{1/2^k} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1}_{1/2^k} \end{bmatrix}, \quad k = 0, ..., m,$$

where $\mathbf{1}_{1/d}$ is the $d \times d$ matrix with $1/d$ as every entry.

- A simple adaptation rule:
  - at each time-step, increase k_n until $\alpha_n = a(k_n)$ satisfies $\mathcal{E}_n^N \geq \tau$
  - overall serial complexity is $O(N)$

Example of adaptive interaction

- assume $N = 2^m$, $m \in \mathbb{N}$ and let

$$\mathbb{A}_N = \left\{ Id, a(1), ..., a(m-1), \mathbf{1}_{1/N} \right\},$$

with

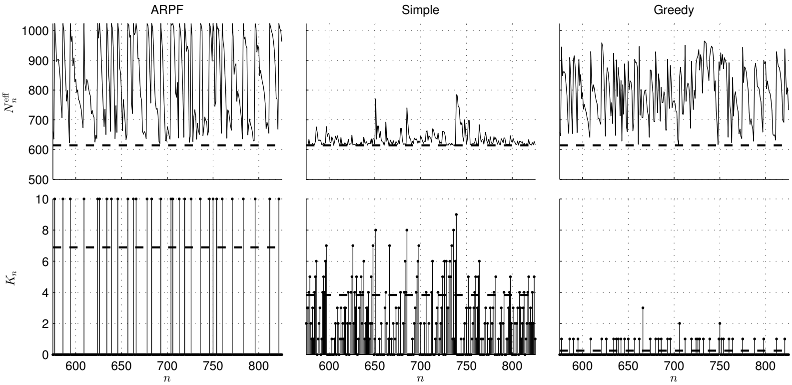
$$a(k) = \begin{bmatrix} \mathbf{1}_{1/2^k} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{1/2^k} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1}_{1/2^k} \end{bmatrix}, \quad k = 0, ..., m,$$

where $\mathbf{1}_{1/d}$ is the $d \times d$ matrix with $1/d$ as every entry.

- A simple adaptation rule:
  - at each time-step, increase k_n until $\alpha_n = a(k_n)$ satisfies $\mathcal{E}_n^N \geq \tau$
  - overall serial complexity is $O(N)$

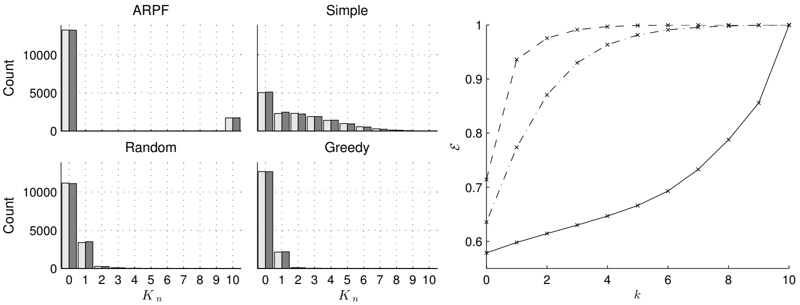
- A greedy adaptation rule:
  - as above, but make pairs, pairs-of pairs, etc. by matching big weights with small weights
  - overall worst case serial complexity is $O(N \log_2 N)$

Numerical illustrations for SV model, $N = 1024$, $\tau = 0.6$



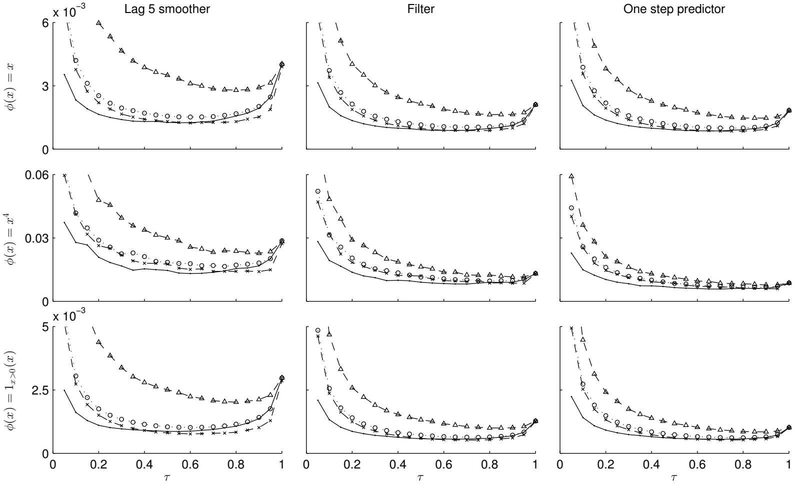
Top: $N_n^{\text{eff}} = N\mathcal{E}_n^N$ vs. time.
Bottom: K_n vs. time. (degree of interaction is 2^{K_n})

Numerical illustrations for SV model, $N = 1024$, $\tau = 0.6$



Left: histograms of K_n over time steps 0 to 15×10^3 and 15×10^3 to 3×10^4
Right: solid=simple, dash-dot=random, dash=greedy

Numerical illustrations for SV model, $N = 1024$



time-averaged mean square filtering errors vs.τ

Right: $\triangle$ =simple, $\circ$ =random, $\times$ =greedy