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

\[
\pi_n(A) := \mathbb{P} \left( X_n \in A \mid Y_{0:n-1} = y_{0:n-1} \right) \\
= \frac{\mathbb{E} \left[ \mathbb{I}[X_n \in A] \prod_{k=0}^{n-1} g(X_k, y_k) \right]}{\mathbb{E} \left[ \prod_{k=0}^{n-1} g(X_k, y_k) \right]} \\
= \frac{\int g(x, y_{n-1}) f(x, A) \pi_{n-1}(dx)}{\int g(x, y_{n-1}) \pi_{n-1}(dx)} = \Phi_n(\pi_{n-1})(A)
\]
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. \( X_0 \sim \pi_0, X_n | X_{n-1} \sim f(X_{n-1}, \cdot) \),

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

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

\[ \text{write } g_n(x) := g(x, y_n) \]

\[ \Rightarrow Z_n := \mathbb{E} \left[ \prod_{k=0}^{n-1} g(X_k, y_k) \right] \]
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 ($\zeta_i^n; n \geq 0), i = 1, \ldots, N$, 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 $W_n^i = W_{n-1}^ig_{n-1}(\zeta_{n-1}^i)$.
    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.
Sequential Importance Sampling

For $n = 0$,

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

Set $W^i_0 = 1$.

Sample $\zeta^i_0 \sim \pi_0$.

For $n \geq 1$,

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

Set $W^i_n = W^i_{n-1} g_{n-1}(\zeta^i_{n-1})$.

Sample $\zeta^i_n \sim f(\zeta^i_{n-1}, \cdot)$.

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

\[
\pi^N_n := \frac{\sum_i W^i_n \delta_{\zeta^i_n}}{\sum_i W^i_n} \approx \frac{\mathbb{E} \left[ \mathbb{I} \left[ X_n \in \cdot \right] \prod_{k=0}^{n-1} g_k(X_k) \right]}{\mathbb{E} \left[ \prod_{k=0}^{n-1} g_k(X_k) \right]} = \pi_n,
\]

\[
Z^N_n := \frac{1}{N} \sum_i W^i_n \approx \mathbb{E} \left[ \prod_{k=0}^{n-1} g_k(X_k) \right] = Z_n.
\]
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 $\zeta_n^i \sim \frac{\sum_j g_{n-1}(\zeta_{n-1}^j)f(\zeta_{n-1}^j, \cdot)}{\sum_j g_{n-1}(\zeta_{n-1}^j)}$.

Recall, $\pi_n(A) = \int g_{n-1}(x) f(x, A) \pi_{n-1}(dx)$, so notice $\int g_{n-1}(x) f(x, A) \pi_{N,n-1}(dx) = \sum_j g_{n-1}(\zeta_{n-1}^j)$.

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

For \( n = 0 \),

For \( i = 1, \ldots, N \),

Sample \( \zeta_i^0 \sim \pi_0 \),

For \( n \geq 1 \),

For \( i = 1, \ldots, N \),

Sample \( \zeta_i^n \sim \frac{\sum_j g_{n-1}(\zeta_j^{n-1})f(\zeta_j^{n-1}, \cdot)}{\sum_j g_{n-1}(\zeta_j^{n-1})} \).

\[
\pi_n^N := \frac{1}{N} \sum_i \delta_{\zeta_i^n} \quad Z_n^N := \prod_{p=0}^{n-1} \left( \frac{1}{N} \sum_i g_p(\zeta_p) \right)
\]
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 $\zeta_n^i \sim \frac{\sum_j g_{n-1}(\zeta_n^j) f(\zeta_n^j, \cdot)}{\sum_j g_{n-1}(\zeta_n^j)}$.

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

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

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 $\varphi : \mathbb{X} \rightarrow \mathbb{R}$ 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 $Z_n^N \rightarrow Z_n$ and $\pi_n^N(\varphi) \rightarrow \pi_n(\varphi)$ as $N \rightarrow \infty$, w.p. 1.
Convergence and time-uniform convergence

For both SIS and BPF,

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

2) for bounded $\varphi : \mathbb{X} \rightarrow \mathbb{R}$ 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 $Z^N_n \rightarrow Z_n$ and $\pi^N_n(\varphi) \rightarrow \pi_n(\varphi)$ as $N \rightarrow \infty$, 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 $p \geq 1$, 

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

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

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

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

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

\[
\sqrt{N} \left[ \pi_n^N(\varphi) - \pi_n(\varphi) \right] \xrightarrow{d} \mathcal{N}(0, \sigma_n^2), \quad \sigma_n^2 = \frac{\mathbb{E} \left[ (\varphi(X_n) - \pi_n(\varphi))^2 \prod_{k=0}^{n-1} g_k(X_k)^2 \right]}{\mathbb{E} \left[ \prod_{k=0}^{n-1} g_k(X_k) \right]^2}.
\]
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 \( \zeta_n^i \sim f(\zeta_{n-1}^i, \cdot) \).

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

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

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

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

\[
\sigma_n^2 = \text{Var}_{\pi_0}(\varphi) \left( \frac{\pi_0(g^2)}{\pi_0(g)^2} \right)^n.
\]
Overview

- Part I: analysis of a “distributed” particle filter
  - *Heine and W.*, ArXiv imminent

- Part II: how much interaction is enough?
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”
- $\theta \in \{1, \ldots, M - 1\}$ is a parameter
- write $y \mod^* x = y - \lfloor (y - 1)/x \rfloor x$.

For $n = 0$,

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

Set $W_{0i} = 1$ and sample $\zeta_0^i \sim \pi_0$

Set $L_i = (i + \theta) \mod^* Mm$,

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

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”
- \( \theta \in \{1, \ldots, M - 1\} \) is a parameter
- write \( y \mod^* x = y - \lfloor (y - 1)/x \rfloor \times x \).

For \( n = 0 \),
  - For \( i = 1, \ldots, N \),
    - Set \( W^i_0 = 1 \) and sample \( \zeta^i_0 \sim \pi_0 \)
    - Set \( L^i = (i + \theta) \mod^* Mm \),
  - 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 \( W^i_n = (Mm)^{-1} \sum_{j \in G_k} W^{L^j}_{n-1} g_{n-1}(\zeta^{L^j}_{n-1}). \)
      - Sample \( \zeta^i_n \sim (W^i_n)^{-1} \sum_{j \in G_k} W^{L^j}_{n-1} g_{n-1}(\zeta^{L^j}_{n-1}) f(\zeta^{L^j}_{n-1}, \cdot) \)

\[
\pi^N_n := \frac{\sum_i W^i_n \delta_{\zeta^i_n}}{\sum_i W^i_n} \quad Z^N_n := \frac{1}{N} \sum_i W^i_n
\]
with $\theta = 0$, LEPF $\equiv$ independent bootstrap particle filters

For $n = 0$,
  For $i = 1, \ldots, N$,
    Set $W_i^0 = 1$ 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 $\zeta_n^i \sim (W_n^i)^{-1} \sum_{j \in G_k} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)$

$\triangleright$ 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 - $\alpha$SMC

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

$\pi_N^n := \sum_i W_n^i \delta_{\zeta_n^i}$
$Z_N^n := \frac{1}{\sum_i W_n^i}$
A general algorithm - $\alpha$SMC

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

For $n = 0$,

- For $i = 1, \ldots, N$,
  - Set $W^i_0 = 1$ and sample $\zeta^i_0 \sim \pi_0$

For $n \geq 1$,

- For $i = 1, \ldots, N$,
  - Set $W^i_n = \sum_j \alpha^{ij} W^j_{n-1} g_{n-1}(\zeta^j_{n-1})$
  - Sample $\zeta^i_n \sim (W^i_n)^{-1} \sum_j \alpha^{ij} W^j_{n-1} g_{n-1}(\zeta^j_{n-1}) f(\zeta^j_{n-1}, \cdot)$

$$
\pi^N_n := \frac{\sum_i W^i_n \delta_{\zeta^i_n}}{\sum_i W^i_n} \quad Z^N_n := \frac{1}{N} \sum_i W^i_n
$$
LEPF and IBPF as $\alpha$SMC

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

$m = M = 3$. left: $\alpha$ for LEPF with $\theta = 1$, right: $\alpha$ for IBPF.
LEPF and IBPF as $\alpha$SMC

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

$M = 3$. left: $\alpha$ for LEPF with $\theta = 1$, right: $\alpha$ for IBPF.
**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 $\varphi : X \to \mathbb{R}$ 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 $\varphi : X \to \mathbb{R}$ and any $M \geq 2$, 

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

with $\sigma_n^2$ depending on $\varphi$, $M$ (and $\theta$ 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

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

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

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

**Thm.** The following hold for the IBPF and LEPF

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

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

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

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

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

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

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” $\alpha$ matrix for the LEPF (for IBPF set $\theta = 0$):

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

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

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

- 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” $\alpha$ matrix for the LEPF (for IBPF set $\theta = 0$):
  $$\alpha_{ij}^\infty := M^{-1} [ (i - 1)/M = (j - \theta - 1)/M ], \quad i, j \in \mathbb{Z}$$

- Let $(I_k, J_k)_{0 \leq k \leq n}$ be the bi-variate, backward Markov chain:
  $$(I_n, J_n) \sim \delta_0 \otimes \delta_0, \quad \mathbb{P}(I_k = i_k, J_k = j_k | I_{k+1} = i_{k+1}, J_{k+1} = j_{k+1}) = \alpha_{i_k+1}^{i_k} \alpha_{j_k+1}^{j_k},$$

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

- Assume $f(x, \cdot) = \pi_0(\cdot)$ and $g_n = g$, then
  $$\sigma_n^2 = \text{Var}_{\pi_0}(\varphi) \mathbb{E}_{0,0} \left[ e^{t Z_n} \right], \quad t := \log \frac{\pi_0(g^2)}{\pi_0(g^2)}.$$
Numerical comparisons

\[ \mathbb{E}_{0,0}[e^{tZ_n}] \text{ vs. } n \text{ for the LEPF} \]

- \( M = 20, \ t = t_0 \)
- \( M = 30, \ t = t_0 \)
- \( M = 30, \ t = .5 \)
- \( M = 40, \ t = .5 \)

\[ R_n \text{ vs. } n \]

- \( M = 20, \ t = t_0 \)
- \( M = 30, \ t = t_0 \)
- \( M = 30, \ t = .5 \)
- \( M = 40, \ t = .5 \)

\[ R_n \text{ and } R_n^N \text{ vs. } n \]

- \( R_n \)
- \( R_{n100} \)
- \( R_{n500} \)
- \( R_{n5000} \)

\[ R_n \text{ vs. } \theta \]

- \( R_{40} \)
- \( R_{70} \)
- \( R_{100} \)
Numerical comparisons

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

\[ R_n^N \text{ vs. } n \]

\[ \text{Approx. error variance vs. } n \]

\[ \varepsilon_n^{Mm} \text{ vs. } n \]

\[ \varepsilon_n^N = \frac{\left(N^{-1} \sum_{i=1}^{N} W_n^i\right)^2}{N^{-1} \sum_{i=1}^{N} (W_n^i)^2} \in [0, 1], \]

\[ \pi_n^N = \frac{\sum_{i=1}^{N} W_n^i \delta \varepsilon_n^i}{\sum_{i=1}^{N} W_n^i}. \]
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 $\alpha_{n-1}$ 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 $\zeta_n^i \sim \frac{\sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) f(\zeta_{n-1}^j, \cdot)}{W_n^i}$

\[
\pi_n^N := \frac{\sum_i W_n^i \delta_{\zeta_n^i}}{\sum_i W_n^i} \quad Z_n^N := \frac{1}{N} \sum_i W_n^i
\]
\( \alpha \text{SMC} - \text{a general algorithm} \)

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 \( W^i_0 = 1 \)
  - Sample \( \zeta^i_0 \sim \pi_0 \)

For \( n \geq 1 \),
- Select \( \alpha_{n-1} \) from \( \mathbb{A}_N \) according to some given functional of \( \{\zeta_0, \ldots, \zeta_{n-1}\} \)
- For \( i = 1, \ldots, N \),
  - Set \( W^i_n = \sum_j \alpha^j_{n-1} W^j_{n-1} g_{n-1}(\zeta^j_{n-1}) \)
  - Sample \( \zeta^i_n \sim \frac{\sum_j \alpha^j_{n-1} W^j_{n-1} g_{n-1}(\zeta^j_{n-1}) f(\zeta^j_{n-1}, \cdot)}{W^i_n} \)

---

\[
\pi^N_n := \frac{\sum_i W^i_n \delta_{\zeta^i_n}}{\sum_i W^i_n} \quad Z^N_n := \frac{1}{N} \sum_i W^i_n
\]

- If \( \mathbb{A}_N = \{Id\} \), \( \alpha \text{SMC} \equiv \text{sequential importance sampling} \)
- If \( \mathbb{A}_N = \{1_{1/N}\} \), \( \alpha \text{SMC} \equiv \text{bootstrap particle filter} \)

(\( 1_{1/N} \) is the matrix with \( 1/N \) as every entry)
Adaptive Resampling PF (Liu and Chen '95) as $\alpha$SMC

- Adaptive Resampling PF is an instance of $\alpha$SMC with $\tau \in (0, 1]$ a threshold, $\Delta_N = \{Id, 1/\sqrt{N}\}$, and

$$\alpha_{n-1} := \begin{cases} 
1/\sqrt{N}, & \text{if } \frac{\left(N^{-1} \sum_i W_{i,n-1} g_{n-1}(\zeta_{i,n-1})\right)^2}{N^{-1} \sum_i \left(W_{i,n-1} g_{n-1}(\zeta_{i,n-1})\right)^2} < \tau. \\
Id, & \text{otherwise.}
\end{cases}$$
Adaptive Resampling PF (Liu and Chen '95) as $\alpha$SMC

- Adaptive Resampling PF is an instance of $\alpha$SMC with $\tau \in (0, 1]$ a threshold, $A_N = \{Id, 1_{1/N}\}$, and

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

- For general $\alpha$SMC, introduce the key quantity:

$$E_N^n := \frac{(N^{-1} \sum_i W_n^i)^2}{N^{-1} \sum_i (W_n^i)^2} = \frac{(N^{-1} \sum_i \sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j))^2}{N^{-1} \sum_i \left( \sum_j \alpha_{n-1}^{ij} W_{n-1}^j g_{n-1}(\zeta_{n-1}^j) \right)^2} \in [0, 1]$$
Adaptive Resampling PF (Liu and Chen ’95) as $\alpha$SMC

- Adaptive Resampling PF is an instance of $\alpha$SMC with $\tau \in (0, 1]$ a threshold, $A_N = \{Id, 1_{1/N}\}$, and

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

- For general $\alpha$SMC, introduce the key quantity:

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

- Notice

\[ \alpha_{n-1} = 1_{1/N} \Rightarrow E^N_n = 1, \]

\[ \alpha_{n-1} = Id \Rightarrow E^N_n = \frac{(N^{-1} \sum_i W_{n-1}^i g_{n-1}(\zeta_{n-1}^i))^2}{N^{-1} \sum_i \left(W_{n-1}^i g_{n-1}(\zeta_{n-1}^i)\right)^2}. \]
Adaptive Resampling PF (Liu and Chen ’95) as $\alpha$SMC

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

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

- For general $\alpha$SMC, introduce the key quantity:

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

- Notice

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

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

- So ARPF can be viewed as enforcing:

$$\inf_n \mathcal{E}_n^N \geq \tau$$
Convergence of $\alpha$SMC

(A) - the entries of the $A_N$-valued random matrix $\alpha_{n-1}$ are measurable w.r.t. $F_{n-1} = \sigma(\zeta_0, ..., \zeta_{n-1})$

(B) - every member of $A_N$ is doubly stochastic
Convergence of $\alpha$SMC

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

(B) - every member of $A_N$ 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 $\varphi : X \rightarrow \mathbb{R}$ 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 $Z^N_n \rightarrow Z_n$ and $\pi^N_n(\varphi) \rightarrow \pi_n(\varphi)$ as $N \rightarrow \infty$, w.p. 1.
Uniform-convergence of \( \alpha \text{SMC} \)

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

\[
f(x, \cdot) \leq \epsilon f(x', \cdot), \quad \forall x, x', \quad \text{and} \quad \sup_{n \geq 0} \sup_{x, x'} g_n(x) \leq \delta.
\]
Uniform-convergence of \( \alpha \text{SMC} \)

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

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

**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 $\alpha$SMC

(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 $\varphi : X \to \mathbb{R}$

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

**Cor.**

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

<table>
<thead>
<tr>
<th>$A_N$</th>
<th>SIS</th>
<th>PF</th>
<th>Adaptive Resampling PF</th>
<th>new instances of $\alpha$SMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>${1d}$</td>
<td>${1_{1/N}}$</td>
<td>${1d, 1_{1/N}}$</td>
<td>a matter of design</td>
<td></td>
</tr>
<tr>
<td>$\alpha_n$ chosen from $A_N$ according to some function of ${\zeta_0, \ldots, \zeta_n}$</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>easy to verify</td>
</tr>
<tr>
<td>$A_N$ contains only doubly stochastic matrices</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>easy to verify</td>
</tr>
<tr>
<td>$\inf_{n \geq 0} E_n^N &gt; 0$ enforced</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>a matter of design</td>
</tr>
</tbody>
</table>
Example of adaptive interaction

\[ A_N = \{ l\mathbf{1}, a(1), \ldots, a(m - 1), \mathbf{1}_{1/N} \} , \]

with

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

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

\[
\mathbb{A}_N = \{ ld, a(1), \ldots, a(m-1), \mathbf{1}_{1/N} \},
\]

with

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

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

\[ A_N = \{ ld, a(1), ..., a(m-1), \mathbf{1}_{1/N} \} , \]

with

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

where $1_{1/d}$ 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: $N_n^{\text{eff}} = N\varepsilon_n^N$ 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 $15 \times 10^3$ and $15 \times 10^3$ to $3 \times 10^4$
Right: solid=simple, dash-dot=random, dash=greedy
Numerical illustrations for SV model, $N = 1024$

Time-averaged mean square filtering errors vs. $\tau$
Right: $\triangle =$simple, $\circ =$random, $\times =$greedy