Variance estimation in the particle filter

Anthony Lee
University of Bristol & Alan Turing Institute

Joint work with Nick Whiteley (University of Bristol)

Biometrika 105(3), 2018

Bayesian Computation for High-Dimensional Statistical Models
Institute for Mathematical Sciences
National University of Singapore
August 30, 2018
Prince George’s Park in the 80s
Outline

Sequential Monte Carlo / particle filters

Variance estimators via ancestral information

Some examples

Some of the theory
Outline

Sequential Monte Carlo / particle filters

Variance estimators via ancestral information

Some examples

Some of the theory
Monte Carlo: approximate sums with random variables.

Notation: $X$ countable, $f : X \rightarrow \mathbb{R}$ and $\mu : X \rightarrow \mathbb{R}_+$,

$$
\mu(f) := \sum_{x \in X} f(x) \mu(x).
$$

If $\mu$ is a distribution, $\mu(1) = 1$ and

$$
\mu(f) = \mathbb{E}[f(X)], \quad X \sim \mu.
$$

Monte Carlo approximation:

$$
\mu^N(f) := \frac{1}{N} \sum_{i=1}^{N} f(\zeta_i), \quad \zeta_i \overset{iid}{\sim} \mu,
$$

satisfies $\mu^N(f) \xrightarrow{a.s.} \mu(f)$ as $N \rightarrow \infty$. 
Sequential Monte Carlo: distributions of interest

- Ingredients:
  1. Initial distribution $\mu$,
  2. Markov transition functions $M_1, \ldots, M_n$,
  3. Non-negative potential functions $G_0, \ldots, G_{n-1}$.
- Define $\gamma_0 := \mu$ and with $\cdot$ denoting pointwise product
  
  \[
  \gamma_p := (\gamma_{p-1} \cdot G_{p-1}) M_p, \quad p \geq 1.
  \]

- Define distributions
  \[
  \eta_p := \frac{\gamma_p}{\gamma_p(1)}, \quad p \geq 0.
  \]

- Equivalently, $\eta_0 = \mu$ and
  \[
  \eta_p = \frac{\eta_{p-1} \cdot G_{p-1}}{\eta_{p-1}(G_{p-1})} M_p = \Phi_p(\eta_{p-1}), \quad p \geq 1.
  \]
Example application: hidden Markov models

- Let $X_0, \ldots, X_n$ be a Markov chain with initial distribution $\mu$ and transitions $M_1, \ldots, M_n$.
- Let $Y_0, \ldots, Y_{n-1}$ be conditionally independent such that $P(Y_p = y_p | X_0:n = x_0:n) = M_p^Y(x_p, y_p)$.

Figure: A graphical model of a hidden Markov model
Hidden Markov model updates

- Let $y_{0:n-1}$ be the observation record, and write
  \[ G_p(x) = \mathcal{M}^y_p(x, y_p), \quad p \in \{0, \ldots, n-1\}. \]
- We have $\gamma_0(x) = \eta_0(x) = \mu(x) = P(X_0 = x)$.
- Then one has
  \[ \gamma_p(x) = P(X_p = x, Y_{0:p-1} = y_{0:p-1}), \quad p \in \{1, \ldots, n\}, \]
  and so
  \[ \eta_p(x) = P(X_p = x \mid Y_{0:p-1} = y_{0:p-1}), \quad p \in \{1, \ldots, n\}. \]
- We see that $Z_n := \gamma_n(1) = P(Y_{0:n-1} = y_{0:n-1})$. 
The forward algorithm (finite state space)

- Initialize: \( \eta_0 \leftarrow \mu \).
- For \( p = 1, \ldots, n \): set
  \[
  \eta_p \leftarrow \frac{(\eta_{p-1} \cdot G_{p-1}) \cdot M_p}{\eta_{p-1}(G_{p-1})}.
  \]
- Set
  \[
  Z_n \leftarrow \prod_{p=0}^{n-1} \eta_p(G_p) = P(Y_{0:n-1} = y_{0:n-1}).
  \]
- If \(|X|\) is very large, or infinite, this is too expensive.
Sequential Monte Carlo [Gordon et al., 1993]

- Target: $\gamma_n \propto \eta_n$ with $Z_n = \gamma_n(1)$.
- We have

$$\eta_p = \Phi_p(\eta_{p-1}), \quad p \in \{1, \ldots, n\}.$$

- Algorithm: Construct $\eta_p^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\zeta_i^p}, \quad p \in \{0, \ldots, n\}$, where

$$\zeta_0 \sim \eta_0, \quad \zeta_i \sim \Phi_p(\eta_{p-1}^N), \quad p \in \{1, \ldots, n\}.$$

- An r.v. $Z_n^N$ is also produced, and with $\gamma_n^N := Z_n^N \eta_n^N$,

$$\gamma_n^N(\varphi) = Z_n^N \frac{1}{N} \sum_{i=1}^{N} \varphi(\zeta_i^n), \quad \mathbb{E}[\gamma_n^N(\varphi)] = \gamma_n(\varphi).$$
A bit more detail

• The transports have the form:

\[ \eta_p = \Phi_p(\eta_{p-1}) = \frac{\eta_{p-1} \cdot G_{p-1}}{\eta_{p-1}(G_{p-1})} M_p. \]

• We instead sample

\[ \zeta^i_p \overset{iid}{\sim} \Phi_p(\eta^N_{p-1}) = \frac{\sum_{j=1}^{N} G_{p-1}(\zeta^j_p) M_p(\zeta^j_p, \cdot)}{\sum_{j=1}^{N} G_{p-1}(\zeta^j_p)}. \]

and define \( \eta^N_p = \frac{1}{N} \sum_{i=1}^{N} \delta_{\zeta^i_p}. \)

• We define

\[ Z^N_n = \prod_{p=0}^{n-1} \eta^N_p(G_p), \quad \text{mirroring} \quad Z_n = \prod_{p=0}^{n-1} \eta_p(G_p). \]
Convergence & asymptotic variance [Del Moral, 2004]

- Under weak conditions, we have
  \[ \gamma_N^n(\varphi) \xrightarrow{a.s.} \gamma_n(\varphi), \quad \eta_N^n(\varphi) \xrightarrow{a.s.} \eta_n(\varphi). \]

- In fact, under still quite weak conditions,
  \[ \sigma^2_n(\varphi) := \lim_{N \to \infty} N \text{var} \left[ \frac{\gamma_N^n(\varphi)}{\gamma_n(1)} \right] = \sum_{p=0}^{n} v_{p,n}(\varphi) < \infty \]
  and we have
  \[ \lim_{N \to \infty} N \mathbb{E} \left[ \left| \eta_N^n(\varphi) - \eta_n(\varphi) \right|^2 \right] = \sigma^2_n(\varphi - \eta_n(\varphi)). \]
Approximating general expectations [Del Moral et al., 2006]

- Ingredients: initial distribution $\mu$, sequence of unnormalized distributions $\mu = \nu_0, \ldots, \nu_n \propto \pi$.

- Let $M_1, \ldots, M_n$ be a sequence of Markov transitions, $G_0, \ldots, G_{n-1}$ a sequence of functions satisfying

$$
\nu_p M_p = \nu_p, \quad G_{p-1} = \nu_p / \nu_{p-1}, \quad p \in \{1, \ldots, n\}.
$$

- E.g. $M_p$ is a $\nu_p$-invariant Metropolis–Hastings transition.

- Then the same setup gives $\gamma_p = \nu_p$ for all $p$,

$$
\eta_n^N(\varphi) \xrightarrow{a.s.} \eta_n(\varphi) = \pi(\varphi),
$$

and

$$
\gamma_n^N(1) \xrightarrow{a.s.} \gamma_n(1) = \nu_n(1).
$$
What about estimating the variance?

- Despite heavy use for over 20 years, until recently the only way to estimate the variance involved running multiple particle filters — impractical.

- We provide an estimate of the variance as a by-product of running a single particle filter, extending Chan and Lai [2013]'s breakthrough contribution.

- We can also estimate individual asymptotic variance terms, \( \nu_{p,n}(\varphi) \) and hence \( \sigma_n^2(\varphi) \).
Outline

Sequential Monte Carlo / particle filters

Variance estimators via ancestral information

Some examples

Some of the theory
Ancestors

- The main / only step is sampling

\[ \zeta_i \overset{iid}{\sim} \Phi_p(\eta_{p-1}^N) = \frac{\sum_{j=1}^{N} G_{p-1}(\zeta_{p-1}^j) M_p(\zeta_{p-1}^j, \cdot)}{\sum_{j=1}^{N} G_{p-1}(\zeta_{p-1}^j)}, \]

so as to construct \( \eta_p^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\zeta_i} \).

- We can look at this as two-stage process: first sample i.i.d.

\[ A_{p-1}^i \sim \text{Categorical} \left( \frac{G_{p-1}(\zeta_{p-1}^1)}{\sum_{j=1}^{N} G_{p-1}(\zeta_{p-1}^j)}, \ldots, \frac{G_{p-1}(\zeta_{p-1}^N)}{\sum_{j=1}^{N} G_{p-1}(\zeta_{p-1}^j)} \right), \]

then \( \zeta_i \overset{ind}{\sim} M_p(\zeta_{p-1}^{A_{p-1}^i}, \cdot). \)

- The random variable \( A_{p-1}^i \) is the index of the ancestor of \( \zeta_p^i \).
Ancestral lineages and eve indices

Figure: A particle system with $n = 3$ and $N = 4$. An arrow from $\zeta_{p-1}^i$ to $\zeta_p^j$ indicates that the ancestor of $\zeta_p^j$ is $\zeta_{p-1}^i$, i.e. $A_{p-1}^j = i$.

- Ancestral lineage: trace ancestor indices backwards from $\zeta_n^i$.
- Eve index: $E_p^i$ is the time 0 index of the ancestor of $\zeta_p^i$. 

First variance estimators

**Theorem.** Let

\[ V_n^N(\varphi) := \eta_n^N(\varphi)^2 - \left( \frac{N}{N-1} \right)^n \frac{1}{N(N-1)} \sum_{i,j:E_n^i \neq E_n^j} \varphi(\zeta_n^i)\varphi(\zeta_n^j). \]

Then the following hold, for bounded \( G_0, \ldots, G_{n-1}, \varphi, \)

1. \( E \{ \gamma_n^N(1)^2 V_n^N(\varphi) \} = \text{var} \{ \gamma_n^N(\varphi) \} \) for all \( N \geq 2, \)
2. \( N V_n^N(\varphi) \to^P \sigma_n^2(\varphi), \)
3. \( N V_n^N(\varphi - \eta_n^N(\varphi)) \to^P \sigma^2(\varphi - \eta_n(\varphi)). \)

- The computational cost is essentially free, one just keeps track of the Eve indices through time.
- \( V_n^N(\varphi) \) can be re-expressed so that it is more obviously \( O(N). \)
- The estimator \( N V_n^N(\varphi - \eta_n^N(\varphi)) \) is \( \left( \frac{N}{N-1} \right)^n \) times the estimator of Chan & Lai (2013) for the limiting CLT variance of \( \eta_n^N(\varphi). \)
Term-by-term asymptotic variance estimators

**Theorem.** For some computable (details later) $v_{0,n}^N(\varphi), \ldots, v_{n,n}^N(\varphi)$, the following hold, for bounded $G_0, \ldots, G_{n-1}, \varphi$,

1. $E \{ \gamma_n^N(1)^2 v_{p,n}^N(\varphi) \} = \gamma_n(1)^2 v_{p,n}(\varphi)$ for all $N \geq 2$,
2. $v_{p,n}^N(\varphi) \xrightarrow{P} v_{p,n}(\varphi)$ and $v_{p,n}^N(\varphi - \eta_n^N(\varphi)) \xrightarrow{P} v_{p,n}(\varphi - \eta_n(\varphi))$.

Hence,

$$\sum_{p=0}^n v_{p,n}^N(\varphi) \xrightarrow{P} \sigma_n^2(\varphi), \quad \sum_{p=0}^n v_{p,n}(\varphi - \eta_n^N(\varphi)) \xrightarrow{P} \sigma_n^2(\varphi - \eta_n(\varphi)).$$

Space complexity $\mathcal{O}(Nn)$: requires storing ancestral indices.
Time complexity $\mathcal{O}(Nn)$: same as running the particle filter.
Outline

Sequential Monte Carlo / particle filters

Variance estimators via ancestral information

Some examples

Some of the theory
Notation: updated measures / estimators

- In the following, \( \hat{\eta}_n = \eta_n \cdot G_n / \eta_n(G_n) \) is an “updated” measure.
  - Think of predictive vs. filtering distributions.

- \( \hat{V}_n(\varphi) \) and \( \hat{v}_{p,n}(\varphi) \) are associated “updated” estimators.
  - There is nothing special really happening here!
**Linear Gaussian example: $N \hat{V}^N(\varphi)$, estimating $\hat{\sigma}^2_n(\varphi)$**

![Graph](image)

(a) $\varphi \equiv 1$

(b) $\varphi = Id - \hat{\pi}^N(Id)$

**Figure:** Estimated asymptotic variances $N \hat{V}^N_n(\varphi)$ (blue dots and error bars for the mean ± one standard deviation) against $\log_2 N$. The red lines correspond to the true asymptotic variances.
Linear Gaussian example: $\hat{v}^N_{p,n}(1), N = 10^5$

Figure: Plot of $\hat{v}^N_{p,n}(\varphi)$ (blue dots and error bars for the mean ± one standard deviation) and $\hat{v}_{p,n}(\varphi)$ (red crosses) at each $p \in \{0, \ldots, n\}$ with $\varphi \equiv 1$. 
Linear Gaussian example: $\hat{v}^N_{p,n}(Id - \hat{\eta}^N_n(Id))$, $N = 10^5$

**Figure:** Plot of $\hat{v}^N_{p,n}(\varphi)$ (blue dots and error bars for the mean ± one standard deviation) and $\hat{v}_{p,n}(\varphi)$ (red crosses) at each $p \in \{0, \ldots, n\}$ with $\varphi = Id - \hat{\pi}^N(Id)$.
SMC sampler example: $NV_n^N(\varphi)$, estimating $\sigma_n^2(\varphi)$

(a) $\varphi \equiv 1$

(b) $\varphi = Id - \pi^N(Id)$

Figure: Estimated asymptotic variances $NV_n^N(\varphi)$ (blue dots and error bars for the mean ± one standard deviation) against $\log_2 N$ for the SMC sampler example.
SMC sampler example: $v_{p,n}^N(\varphi)$, 1 iteration per kernel

Figure: Plot of $v_{p,n}^N(\varphi)$ (blue dots and error bars for the mean ± one standard deviation) at each $p \in \{0, \ldots, n\}$. 

(a) $\varphi \equiv 1$

(b) $\varphi = \text{Id} - \pi(\text{Id})$
SMC sampler example: $v_{p,n}^N(\varphi)$, 10 iterations per kernel

(a) $\varphi \equiv 1$

(b) $\varphi = Id - \pi(Id)$

Figure: Plot of $v_{p,n}^N(\varphi)$ (blue dots and error bars for the mean ± one standard deviation) at each $p \in \{0, \ldots, n\}$. 
What about i.i.d. replicates?

- For fixed $N$, consistent estimation of $\text{var}(\gamma^N(\varphi)/\gamma(1))$ using sample variance and mean of i.i.d. replicates is straightforward.
- Lack-of-bias of $\gamma^N(1)^2 V^N_n(\varphi)$ allows an alternative estimate using replicates of $\gamma^N_n(1)$ and $V^N_n(\varphi)$.

![Figure: Plot of the standard estimate of $\text{var} [\hat{\gamma}^N_n(1)/\hat{\gamma}_n(1)]$ (blue) and the alternative estimate based on $\hat{V}^N_n(1)$ (red) against no. of replicates in the two examples, with $N = 10^3$.](image-url)
Outline

Sequential Monte Carlo / particle filters

Variance estimators via ancestral information

Some examples

Some of the theory
Second moment of $\gamma_n^N(\varphi)$

- Cérou et al. [2011]: for certain measures $\{\mu_b : b \in \{0, 1\}^{n+1}\}$,
  \[
  \mathbb{E} \left[ \gamma_n^N(\varphi)^2 \right] = \sum_{b \in \{0,1\}^{n+1}} \left[ \left( \frac{1}{N} \right)^{|b|} \left( 1 - \frac{1}{N} \right)^{|1-b|} \right] \mu_b(\varphi \otimes 2).
  \]

- These measures have a nice interpretation:
  \[
  \mu_b(\varphi) := \mathbb{E}_b \left[ \varphi(X_n, X'_{n}) \prod_{p=0}^{n-1} G_p(X_p) G_p(X'_p) \right],
  \]
  where $(X_p, X'_p)_{0 \leq p \leq n}$ is a Markov chain defined by $\mu$, $M_1, \ldots, M_n$ and $b$.

- Compare with
  \[
  \gamma_n(\varphi) = \mathbb{E} \left[ \varphi(X_n) \prod_{p=0}^{n-1} G_p(X_p) \right].
  \]
Second moment measures

• For $b \in \{0, 1\}^{n+1}$, $\varphi : X \times X \to \mathbb{R}$

$$
\mu_b(\varphi) := \tilde{E}_b \left[ \varphi(X_n, X'_n) \prod_{p=0}^{n-1} G_p(X_p)G_p(X'_p) \right],
$$

where $(X_p, X'_p) \sim \tilde{M}_p^b(X_{p-1}, X'_{p-1}; \cdot)$.

• When $b_p = 0$,

$$
X_p \sim M_p(X_{p-1}, \cdot), \quad X'_p \sim M_p(X'_{p-1}, \cdot),
$$

independently, and when $b_p = 1$,

$$
X'_p = X_p \sim M_p(X_{p-1}, \cdot).
$$

• When $b = 0$, we obtain $\mu_0(\varphi \otimes^2) = \gamma_n(\varphi)^2$. 
Diagram for $\mu_0$ (top) and $\mu_{e_3}$ (bottom)
Genealogical tracing variables

- Consider the particle system simulated up to time $n$.
- Define auxiliary random variables $K^1 = (K_0^1, \ldots, K_n^1)$ and $K^2 = (K_0^2, \ldots, K_n^2)$, with the following sampling interpretation:

1. $K^1$ is an ancestral lineage: sample $K_n^1$ uniformly from \{1, \ldots, N\}, then for $p = n, \ldots, 1$ set $K_{p-1}^1 = A_{p-1}^{K_p^1}$.
2. $K^2$ consists of possibly “broken” ancestral lineages: sample $K_n^2$ uniformly from \{1, \ldots, N\}, and trace back an ancestral lineage as above, but when a “collision” $K_p^2 = K_p^1$ occurs, sample $K_{p-1}^2$ with probability proportional to $G_{p-1}(\zeta_{p-1}^{k_{p-1}^2})$.

- Let $C(A, \zeta; k^{1:2})$ be the conditional p.m.f. of $K^1, K^2$ given all ancestor indices $A$ and particles $\zeta$ up to time $n$. 
A realization of \((K^1, K^2)\) (red, blue)

\[ k^1 = (4, 4, 3, 1, 2, 3), \quad k^2 = (2, 1, 2, 1, 3, 4). \quad k^{1:2} \text{ related to } e_3. \]
Particle approximations of $\mu_b$

Define, for $b \in \{0, 1\}^{n+1}$, and with $N \geq 2$,

$$
\mu^N_b := \left[ \prod_{p=0}^{n} N^{|b_p|} \left( \frac{N}{N-1} \right)^{|1-b|} \right] \gamma_n N (1)^2 \sum_{k^{1:2} \in \mathcal{I}(b)} C(A, \zeta; k^{1:2}) \delta_{(\zeta^{k_1, k_2}, \zeta^{k_2})},
$$

where $\mathcal{I}(b) := \{ k^{1:2} \in \{1, \ldots, N\}^2 : k^1_p = k^2_p \iff b_p = 1\}$.

**Theorem.** For any $b \in \{0, 1\}^{n+1}$ and bounded $\varphi$,

$$
\mathbb{E} \left[ \mu^N_b(\varphi) \right] = \mu_b(\varphi),
$$

and

$$
\sup_{N \geq 1} \sqrt{N} \mathbb{E} \left[ \left( \mu^N_b(\varphi) - \mu_b(\varphi) \right)^2 \right]^{\frac{1}{2}} < +\infty.
$$
Particle approximations of $\mu_0$

- When $b = 0$, we obtain

$$
\mu_0^N(\varphi \otimes 2) := \gamma_n^N(1)^2 \left( \frac{N}{N - 1} \right)^{n+1} \frac{1}{N^2} \sum_{i,j: E_n^i \neq E_n^j} \varphi(\zeta_n^i)\varphi(\zeta_n^j),
$$

where $E_n^i$ is the index of the time 0 ancestor of $\zeta_n^i$.

- We can compute this in $O(N)$ time since

$$
\frac{1}{N^2} \sum_{i,j: E_n^i \neq E_n^j} \varphi(\zeta_n^i)\varphi(\zeta_n^j) = \eta_n^N(\varphi)^2 - \sum_{i=1}^N \left[ \frac{1}{N} \sum_{j: E_n^j = i} \varphi(\zeta_n^j) \right]^2.
$$

- An unbiased estimator of $\text{var} \left[ \gamma_n^N(\varphi) \right]$ is simply

$$
\gamma_n^N(\varphi)^2 - \mu_0^N(\varphi \otimes 2).
$$
Variance estimator: lack-of-bias

- We use the lack-of-bias:

\[
E \left[ \gamma_n^N(\varphi) \right] = \gamma_n(\varphi).
\]

and

\[
E \left[ \mu_0^N(\varphi^{\otimes 2}) \right] = \mu_0(\varphi^{\otimes 2}) = \gamma_n(\varphi)^2.
\]

- Indeed,

\[
\text{var} \left[ \gamma_n^N(\varphi) \right] = E \left[ \gamma_n^N(\varphi)^2 \right] - E \left[ \gamma_n^N(\varphi) \right]^2
\]
\[
= E \left[ \gamma_n^N(\varphi)^2 \right] - \gamma_n(\varphi)^2
\]
\[
= E \left[ \gamma_n^N(\varphi)^2 \right] - \mu_0(\varphi^{\otimes 2})
\]
\[
= E \left[ \gamma_n^N(\varphi)^2 - \mu_0^N(\varphi^{\otimes 2}) \right].
\]
Variance estimators: consistency

Define

\[ V_N^n(\varphi) := \left[ \gamma_N^n(\varphi)^2 - \mu_0^N(\varphi^\otimes 2) \right] / \gamma_N^n(1)^2 \]

and

\[ v_{p,n}^N(\varphi) := \left[ \mu_{ep}^N(\varphi^\otimes 2) - \mu_0^N(\varphi^\otimes 2) \right] / \gamma_N^n(1)^2. \]

\[ \text{Remark. } NV_0^N(\varphi) \text{ is the unbiased sample variance.} \]

Theorem (again). For any bounded \( \varphi \), as \( N \to \infty \).

1. \( NV_n^N(\varphi) \xrightarrow{P} \sigma_n^2(\varphi) \) and \( NV_n^N(\varphi - \eta_n^N(\varphi)) \xrightarrow{P} \sigma_n^2(\varphi - \eta_n^N(\varphi)) \),

2. \( v_{p,n}^N(\varphi) \xrightarrow{P} v_{p,n}(\varphi) \) and \( v_{p,n}^N(\varphi - \eta_n^N(\varphi)) \xrightarrow{P} v_{p,n}(\varphi - \eta_n(\varphi)) \),

where one has

\[ \sigma_n^2(\varphi) = \lim_{N \to \infty} N \text{ var } \left[ \gamma_n^N(\varphi) / \gamma_n(1) \right] = \sum_{p=0}^{n} v_{p,n}(\varphi). \]
Computational complexity

- Efficient algorithms for computing $V_n^N(\varphi)$ and $\nu_{p,n}^N(\varphi)$ satisfy

| Estimate                        | Time complexity | Space complexity |
|---------------------------------|-----------------|------------------|
| $\gamma_n^N(\varphi)$ or $\eta_n^N(\varphi)$ | $O(Nn)$         | $O(N)$           |
| $V_n^N(\varphi)$                | $O(Nn)$         | $O(N)$           |
| $\nu_{p,n}^N(\varphi)$         | $O(Nn)$         | $O(Nn)$          |

- Calculating $V_n^N(\varphi)$ is $O(N)$ after computing $\gamma_n^N(\varphi)$.
- Calculating $\nu_{p,n}^N(\varphi)$ requires some recursive computations and storage of the genealogies $A_0, \ldots, A_{n-1}$.
- Evaluation only of $\varphi$ and the potentials $(G_p)_{p \geq 0}$ is required, using the output of a single particle filter.
- Computational time is actually negligible.
Final comments

• As $n \to \infty$ with $N$ fixed the particle system degenerates.
• If one is only interested in approximating $\text{var}(\eta_n^N(\varphi))$ one can instead use a fixed-lag based extension of Chan and Lai [2013] proposed by Olsson and Douc [2018].
  - No longer any lack-of-bias results; introduces bias to reduce variance.
• In many situations, $V_n^N(1)$ is a reliable diagnostic when it is less than, say, 0.5.
F. Cérou, P. Del Moral, and A. Guyader. A nonasymptotic theorem for unnormalized Feynman–Kac particle models. *Ann. Inst. H. Poincaré Probab. Statist.*, 47(3):629–649, 2011.

H. P. Chan and T. L. Lai. A general theory of particle filters in hidden Markov models and some applications. *Ann. Statist.*, 41(6):2877–2904, 2013.

P. Del Moral. *Feynman–Kac formulae: Genealogical and Interacting Particle Systems with Applications*. Springer Verlag, New York, 2004.

P. Del Moral, A. Doucet, and A. Jasra. Sequential Monte Carlo samplers. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 68(3):411–436, 2006.

N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *Radar and Signal Processing, IEE Proceedings F*, 140(2):107–113, apr 1993.

A. Lee and N. Whiteley. Variance estimation in the particle filter. *Biometrika*, 105(3):609–625, 2018.

J. Olsson and R. Douc. Numerically stable online estimation of variance in particle filters. *Bernoulli*, 2018. To appear.