Sequential Monte Carlo

some persectives from outside neutron transport

Adam M. Johansen

MathRad Workshop, Bournemouth, May 18-19th, 2023

University of Warwick
a.m.johansen@warwick.ac.uk
https://go.warwick.ac.uk/amjohansen/talks/

Outline

- Background / History
- An Abstract Framework
- Some Improvements
- Some (Useful?) Theory
- Reflections

Background / History

A Very Selective History

- Branching systems date back to the dawn of Monte Carlo (Kahn and Harris, 1951).
- Filtering applications led to popularisation in engineering and statistics (Gordon et al., 1993; Stewart and McCarty Jr, 1992)
- Rigous formulation and analysis “begins” with Del Moral (1995); see Del Moral (2004) and Del Moral (2013).
- Good textbook treatment provided by Chopin and Papaspiliopoulos (2020).
- Recent introductions include Doucet and Johansen (2011) (generic with filtering and smoothing focus); Doucet and Lee (2018); (graphical models); Dai et al. (2022) (SMC samplers).

What is Sequential Monte Carlo?

Sequential Monte Carlo (SMC)

Approximating each of a sequence[1] of distributions using (weighted) empirical distributions of a particle system undergoing mutation and selection dynamics.

$$
\begin{aligned}
\text{SMC} &\approx \text{(mean field) particle approximation (of a Feynman-Kac flow)}\\
&\approx \text{particle filters}\\
&\approx \text{cloning}\\
&\approx \text{genealogical interacting particle systems}\\
&\approx \text{``go-with-the-winner''}\\
&\approx \text{(simple) genetic algorithms}
\end{aligned}
$$

[1] I focus on discrete time / generational algorithms but there are continuous time analogues.

An Abstract Framework

Discrete-time Feynman-Kac Formulæ: a framework

- Ingredients: Markovian dynamics + environment
  - "Initial distribution", μ
  - "Transition kernels", $K_2, K_3, \ldots$
  - "Potential functions", $G_1, G_2, \ldots$
- Describes the law of a particle (X_t) evolving in a potential.
- Typically interested in:
  - Average product of potentials experienced up to time t

$$Z_t := \mathbb{E}\left[\prod_{s=1}^{t} G_s(X_s)\right]$$

  - Law of process *twisted* by those potentials:

$$\eta_t(\varphi) := \mathbb{E}\left[\prod_{s=1}^{t} G_s(X_s)\varphi(X_t)\right] / Z_t$$

A Concrete Example

Single particle moving in an absorbing medium.

- μ — distribution of initial location
- $K_t(x_{t-1}, dx_t)$ — stochastic dynamics over time interval t
- $G_t(x)$ — probability particle is not absorbed at x
- The normalizing constant corresponds to survival probability:

$$Z_t = \mathbb{E}\left[\prod_{s=1}^{t} G_s(X_s)\right] = \int \mu(x_1) G(x_1) \prod_{s=2}^{t} K_s(x_{s-1}, x_s) G_s(x_s) dx_{1:t}$$

- The law of a particle conditional upon its survival is then:

$$\eta_t(A) = \frac{\int_A \mu(x_1) G(x_1) \prod_{s=2}^{t} K_s(x_{s-1}, x_s) G_s(x_s) dx_{1:t}}{\int \mu(x_1) G(x_1) \prod_{s=2}^{t} K_s(x_{s-1}, x_s) G_s(x_s) dx_{1:t}}$$
$$= \frac{\mathbb{P}(X_t \in A, \text{survive to } t)}{\mathbb{P}(\text{survive to } t)}$$

A Simple SMC Algorithm "sequential importance resampling"

SMC / SIR

$t = 1$: initialize

- sample $X_1^1, \ldots, X_1^N \sim \mu$

$t > 1$: iterate

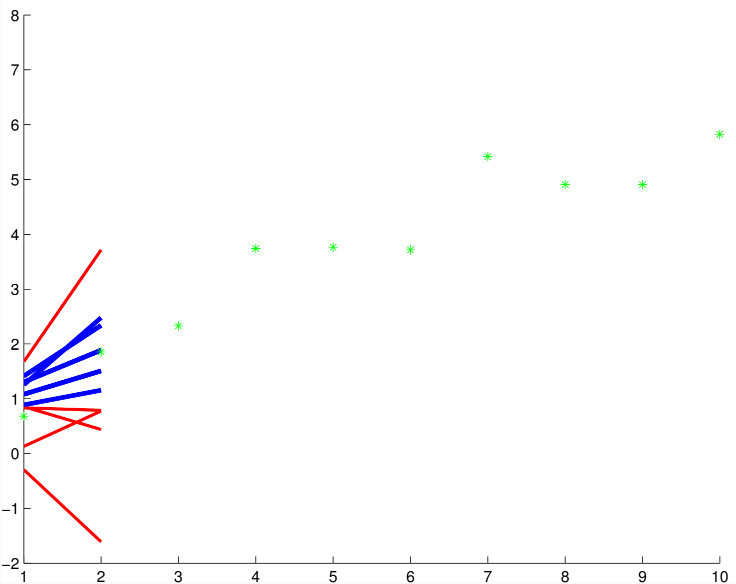
- sample

$$X_t^1, \ldots, X_t^N \sim \frac{\sum_{j=1}^N G_{t-1}(X_{t-1}^j) K_t(X_{t-1}^j, \cdot)}{\sum_{k=1}^N G_{t-1}(X_{t-1}^k)}$$

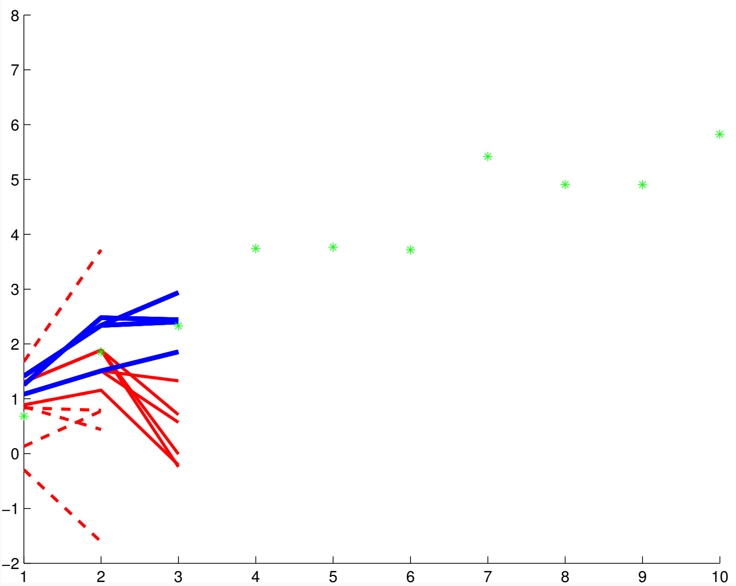
- Approximate $\hat{\eta}_t(dx_t)$, Z_t with

$$\hat{\eta}_t(dx_t) = \frac{\sum_{j=1}^N G_t(X_t^j)\delta_{X_t^j}}{\sum_{k=1}^N G_t(X_t^k)}, \qquad Z_t^N = \prod_{s=1}^t \frac{1}{N} \sum_{j=1}^N G_s(X_s^k)$$

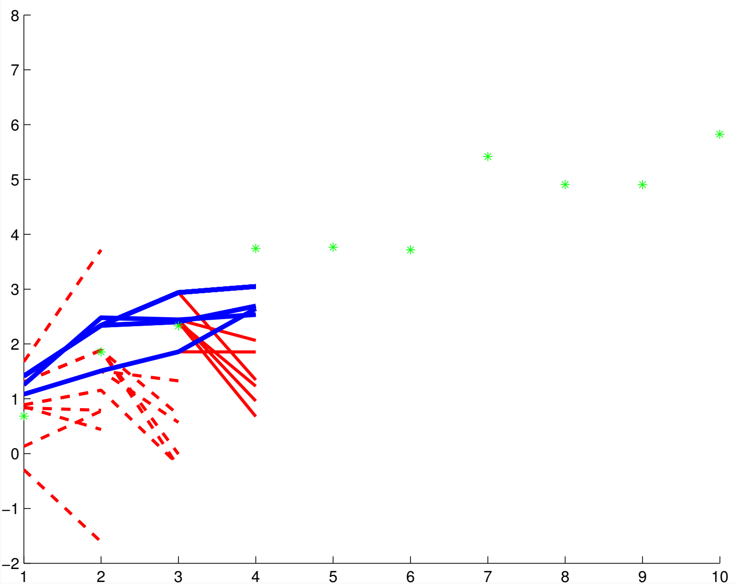
Iteration 2



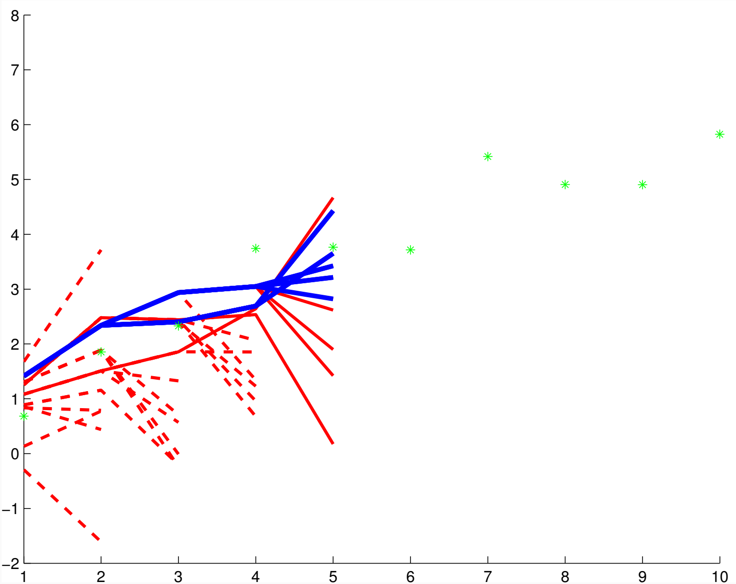
Iteration 3



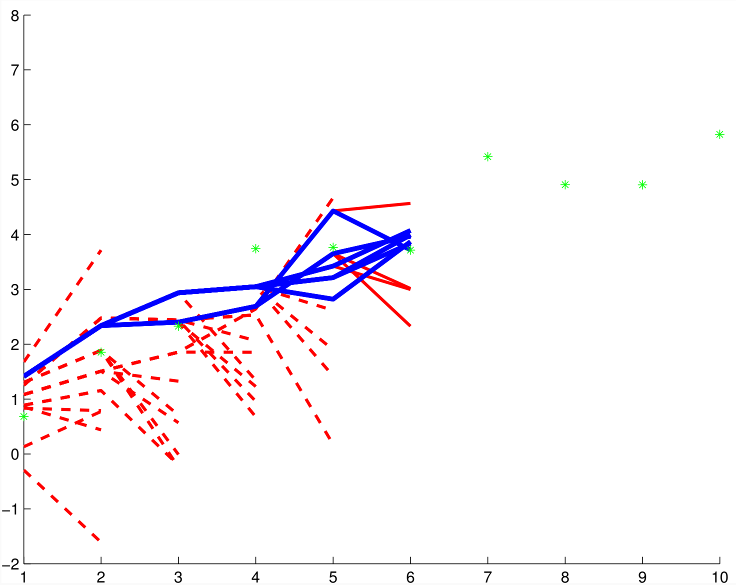
Iteration 4



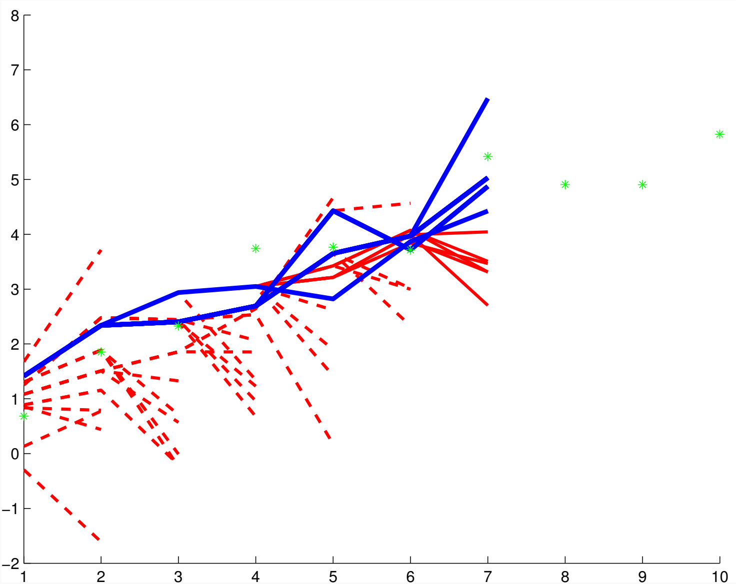
Iteration 5



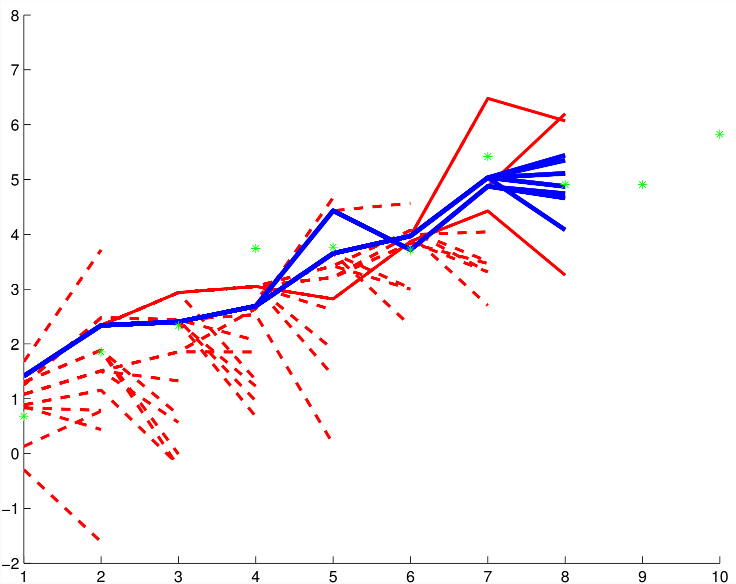
Iteration 6



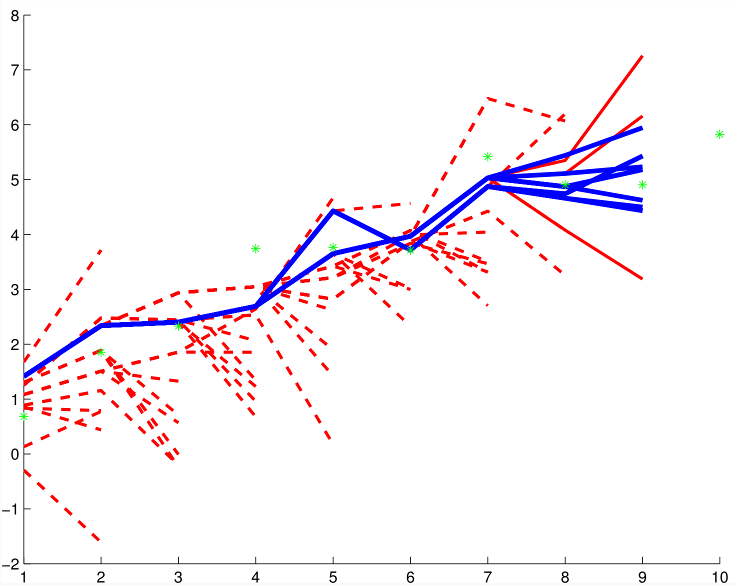
Iteration 7



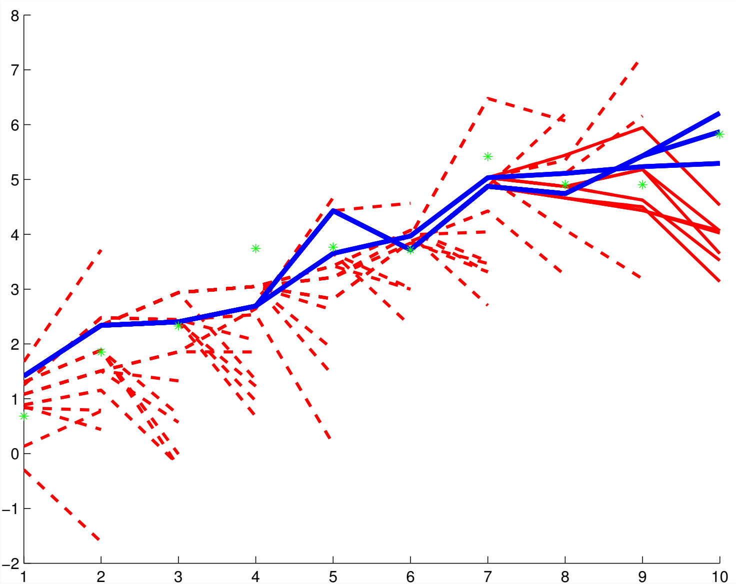
Iteration 8



Iteration 9



Iteration 10



Continuous Time

Require finite representations to store trajectories:

- Exploit specific properties of trajectories
  - Tractable transition density (e.g. discretely oberved OU processes)
  - Pure jump processes (e.g., cloning; Angeli et al. (2021))
  - Exact simulation (e.g. filtering; Fearnhead et al. (2008))
  - ϵ-strong simulation (e.g. rare event simulation; Hodgson et al. (2022))
- Discretise time and apply discrete-time algorithm
- Develop apparently continuous-time algorithms but implement discretizations — see Del Moral and Miclo (2000).

Some Improvements

Better Selection Mechanisms

Generational whole population selection:

- Better *resampling schemes*, see Gerber et al. (2019): residual, stratified, systematic(?), sorted versions, tree-based branching approximation, . . .
- Adaptive resampling

Localized selection events (birth/death; natural in continuous time):

- Continuous time variants — see Rousset (2006)
- Cloning mechanisms — see Angeli et al. (2021)

Better “proposal distributions”

- *Strutural stability:* If $K_t(x_{t-1}, dx_t)G_t(x_t) = \tilde{K}_t(x_{t-1}, dx_t)\tilde{G}_t(x_t)$ for every t then $\{K_t, G_t\}$ and $\{\tilde{K}_t, G_t\}$ define essentially the same model.
- *Locally optimal proposals* (Doucet et al., 2000): choose $\tilde{K}_t \propto G_t \cdot K_t$; leads to

$$\text{“}\tilde{G}_t(x_t) = \int K_t(x_{t-1}, dx_t)G_t(x_t)\text{”}$$

  but this is easily made rigorous[2].
- *Auxiliary particle methods* (Pitt and Shephard, 1999; Johansen and Doucet, 2008).
- *Marginal particle methods* (Klass et al., 2005; Crucinio and Johansen, 2023).

[2] Spatial extension; or slight model redefinition — in reality, essential.

Changing the Model/Algorithm — Twisting

What about the future?

Information propagates through time...

- Twisted particle methods (Whiteley and Lee, 2014).
- Lookahead methods (incl. piloting, stochastic piloting; cf. Lin et al. (2013)).
- Block-sampling methods (Doucet et al., 2006).
- Controlled-methods (Guarniero et al., 2017; Heng et al., 2020)

Path and Parameter Estimation

Improving trajectory estimates (see, e.g. Briers et al. (2010)):

- Forward-backward algorithms
- Backward-information filters
- Fixed-lag methods

Estimating static parameters (cf. Kantas et al. (2015)):

- Stochastic gradient methods
- Particle MCMC (Andrieu et al., 2010); SMC^2 (Chopin et al., 2013)

Some (Useful?) Theory

What can be shown

A version of the fist six can be extracted from Del Moral (2004).

- Moment bounds on errors.
- Strong (and weak) Laws of Large Numbers.
- Weak convergence of measures.
- Central Limit Theorems (and Berry-Esseen, Donsker, functional forms. . .)
- Propagation of chaos.
- Bias bounds.
- Variance can be estimated from a *single run* (Lee and Whiteley, 2018)

Two main techniques: dynamic semigroup methods and explicit recursive formulations (e.g. Crisan and Doucet (2002); Chopin (2004); Douc and Moulines (2008)).

Reflections

Reflections

- Many of these things has close analogues in the neutron transport setting.
- There may be significant opportunities for both communities to benefit from better communication.
- Some open methodological problems include:
    - How to effective perform online filtering, smoothing and parameter estimation in high-dimensional hidden Markov Models.
    - How best to perform online parameter estimation in statistical settings.
    - How to leverage the power of block-sampling or controlled methods in greater generality.
    - How to avoid discretisation in greater generality.

References

C. Andrieu, A. Doucet, and R. Holenstein. Particle Markov chain Monte Carlo. *Journal of the Royal Statistical Society B*, 72(3):269–342, 2010.

L. Angeli, S. Grosskinsky, and A. M. Johansen. Limit theorems for cloning algorithms. *Stochastic Processes and their Applications*, 138:117–152, 2021. doi: 10.1016/j.spa.2021.04.007.

M. Briers, A. Doucet, and S. Maskell. Smoothing algorithms for state space models. *Annals of the Institute of Statistical Mathematics*, 62(1):61–89, 2010.

N. Chopin. Central limit theorem for sequential Monte Carlo methods and its applications to Bayesian inference. *Annals of Statistics*, 32(6):2385–2411, December 2004.

N. Chopin and O. Papaspiliopoulos. *An introduction to sequential Monte Carlo*. Springer, 2020.

N. Chopin, P. Jacob, and O. Papaspiliopoulos. SMC^2: an efficient algorithm for sequential analysis of state space models. *Journal of the Royal Statistical Society B*, 75(3):397–426, 2013.

D. Crisan and A. Doucet. A survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on Signal Processing*, 50(3):736–746, March 2002.

F. R. Crucinio and A. M. Johansen. A divide-and-conquer sequential Monte Carlo approach to high dimensional filtering. *Statistica Sinica*, 2023. doi: 10.5705/ss.202022.0243. In press.

C. Dai, J. Heng, P. E. Jacob, and N. Whiteley. An invitation to sequential Monte Carlo samplers. *Journal of the American Statistical Association*, 117(539):1587–1600, 2022. doi: 10.1080/01621459.2022.2087659.

P. Del Moral. Nonlinear filtering using random particles. *Theory of Probability and Its Applications*, pages 690–701, 1995.

P. Del Moral. *Feynman-Kac formulae: genealogical and interacting particle systems with applications*. Probability and Its Applications. Springer Verlag, New York, 2004.

P. Del Moral. *Mean Field Integration*. Chapman Hall, 2013.

P. Del Moral and L. Miclo. Branching and interacting particle systems and approximations of Feynman-Kac formulæ with applications to non-linear filtering. In J. Azéma, M. Emery, M. Ledoux, and M. Yor, editors, *Séminaire de Probabilités XXXIV*, volume 1729 of *Lecture Notes in Mathematics*, pages 1–145. Springer Verlag, Berlin, 2000.

R. Douc and E. Moulines. Limit theorems for weighted samples with applications to sequential monte carlo methods. *Annals of Statistics*, 36(5):2344–2376, 2008.

A. Doucet and A. M. Johansen. A tutorial on particle filtering and smoothing: Fiteen years later. In D. Crisan and B. Rozovsky, editors, *The Oxford Handbook of Nonlinear Filtering*, pages 656–704. Oxford University Press, 2011.

A. Doucet and A. Lee. Sequential Monte Carlo methods. In M. Maathuis, M. Drton, S. L. Lauritzen, and M. Wainwright, editors, *Handbook of Graphical Models*, pages 165–189. CRC Press, 2018.

A. Doucet, S. Godsill, and C. Andrieu. On sequential simulation-based methods for Bayesian filtering. *Statistics and Computing*, 10(3):197–208, 2000.

A. Doucet, M. Briers, and S. Sénécal. Efficient block sampling strategies for sequential Monte Carlo methods. *Journal of Computational and Graphical Statistics*, 15(3):693–711, 2006.

P. Fearnhead, O. Papaspiliopoulos, and G. O. Roberts. Particle filters for partially-observed diffusion. *Journal of the Royal Statistical Society B*, 70(4):755–777, 2008.

M. Gerber, N. Chopin, and N. Whiteley. Negative association, ordering and convergence of resampling methods. *Annals of Statistics*, 47(4):2236–2260, 2019.

N. J. Gordon, S. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings-F*, 140(2):107–113, April 1993. doi: 10.1049/ip-f-2.1993.0015.

P. Guarniero, A. M. Johansen, and A. Lee. The iterated auxiliary particle filter. *Journal of the American Statistical Association*, 112(520):1636–1647, 2017. doi: 10.1080/01621459.2016.1222291.

J. Heng, A. N. Bishop, G. Deligiannidis, A. Doucet, et al. Controlled sequential Monte Carlo. *Annals of Statistics*, 48(5):2904–2929, 2020.

J. Hodgson, A. M. Johansen, and M. Pollock. Unbiased simulation of rare events in continuous time. *Methodology and Computing in Applied Probability*, 24:2123–2148, 2022. doi: 10.1007/s11009-021-09886-2.

A. M. Johansen and A. Doucet. A note on the auxiliary particle filter. *Statistics and Probability Letters*, 78(12): 1498–1504, September 2008. doi: 10.1016/j.spl.2008.01.032.

H. Kahn and T. E. Harris. Estimation of particle transmission by random sampling. *National Bureau of Standards Applied Mathematics Series*, 12:27–30, 1951.

N. Kantas, A. Doucet, S. S. Singh, J. M. Maciejowski, and N. Chopin. On particle methods for parameter estimation in general state-space models. *Statistical Science*, 30(3):328–351, 2015.

M. Klass, N. de Freitas, and A. Doucet. Towards practical n^2 Monte Carlo: The marginal particle filter. In *Proceedings of Uncertainty in Artificial Intelligence*, 2005.

A. Lee and N. Whiteley. Variance estimation in the particle filter. *Biometrika*, 105(1):609–625, 2018.

M. Lin, R. Chen, and J. S. Liu. Lookahead strategies for sequential Monte Carlo. *Statistical Science*, 28(1): 69–94, 2013.

M. K. Pitt and N. Shephard. Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association*, 94(446):590–599, 1999.

M. Rousset. On the control of an interacting particle estimation of Schrödinger ground states. *SIAM journal on mathematical analysis*, 38(3):824–844, 2006.

L. Stewart and P. McCarty Jr. Use of Bayesian belief networks to fuse continuous and discrete information for target recognition, tracking, and situation assessment. In *Signal Processing, Sensor Fusion, and Target Recognition*, volume 1699, pages 177–185. SPIE, 1992. doi: 10.1117/12.138224.

N. Whiteley and A. Lee. Twisted particle filters. *Annals of Statistics*, 42(1):115–141, 2014.