Milestones in Probability and Statistics
The Bootstrap (Particle) Filter

Adam M. Johansen

October 4th, 2024

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

Warwick Statistics

Our Starting Point

Gordon et al. (1993): 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. [11029 citations]

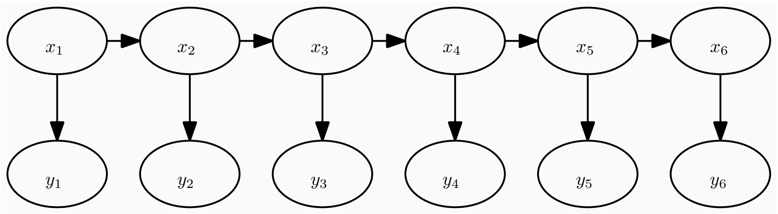
See also:

- Stewart and McCarty Jr (1992): contains essentially the same idea without spelling out the algorithm. [46 citations]
- Various technical reports by Pierre Del Moral and co-authors from 1992-93. [44 citations]

(Citation data from google scholar as of 29th September, 2023.)

This talk aims to give a (inevitably rather selective; lots of interesting contibutions have been omitted) survey of the way this ideas has evolved through and influenced the wider literature.

(General State Space) Hidden Markov Models



Prior $p(x_1) = \mu(x_1)$

Dynamics $p(x_t|x_{1:t-1}, y_{1:t-1}) = f(x_t|x_{t-1})$

Observation $p(y_t|x_{1:t}, y_{1:t-1}) = g(y_t|x_t)$

Filtering $p(x_t|y_{1:t})$

Example Tracking a moving target, e.g. air traffic control.

The Filtering Recursions

Joint

$$
\begin{aligned}
p(x_{1:t}|y_{1:t}) \propto & p(x_{1:t}|y_{1:t-1})p(y_t|x_{1:t}, y_{1:t-1}) \\
= & p(x_{1:t-1}|y_{1:t-1})p(x_t|x_{1:t-1}, y_{1:t-1})p(y_t|x_{1:t}, y_{1:t-1}) \\
= & p(x_{1:t-1}|y_{1:t-1})f(x_t|x_{t-1})g(y_t|x_t)
\end{aligned}
$$

Marginal

$$
\begin{aligned}
p(x_t|y_{1:t}) \propto & p(x_t|y_{1:t-1})p(y_t|x_t, y_{1:t-1}) \\
= & \int p(x_{t-1}|y_{1:t-1})p(x_t|x_{t-1}, y_{1:t-1})dx_{t-1}p(y_t|x_t, y_{1:t-1}) \\
= & \int p(x_{t-1}|y_{1:t-1})f(x_t|x_{t-1})dx_{t-1}g(y_t|x_t)
\end{aligned}
$$

Bayesian bootstrap filter—original(ish) presentation

Suppose we have a set of random samples $\{x_{t-1}(i) : i = 1, \ldots, N\}$ from the PDF $p(x_{t-1}|y_{1:t-1})$:

prediction each sample is passed through the system model

$$x_t^*(i) \sim f(x_t|x_{t-1}^{(i)})$$

update obtain a normalized weight for each sample

$$w_t^i = \frac{g(y_t|x_t^{*(i)})}{\sum_{j=1}^{i} g(y_t|x_t^{*(j)})}$$

sample N times from the discrete distribution placing mass w_i on point $x_t^{*(i)}$ for each i to obtain $\{x_{t-1}(i) : i = 1, \ldots, N\}$

"enhancements" **roughening** add some noise after resampling

prior editing throw away particles inconsistent with next observation

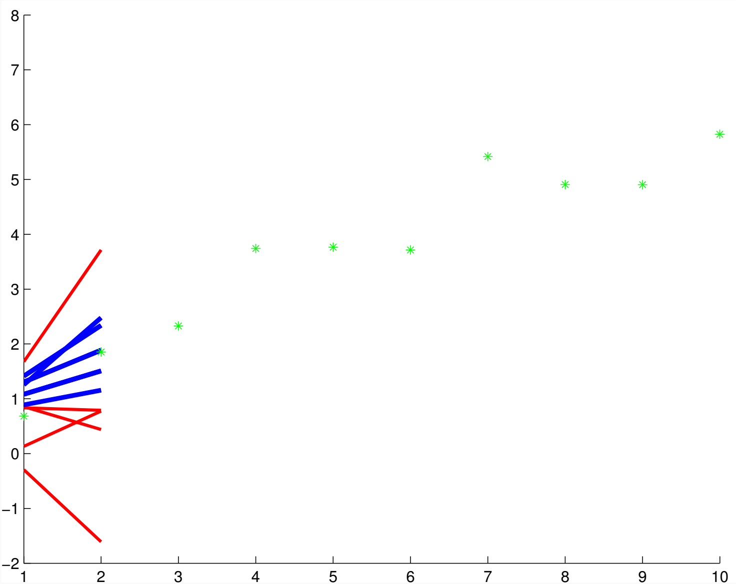
A Toy Example

Follow a signal measured with noise.

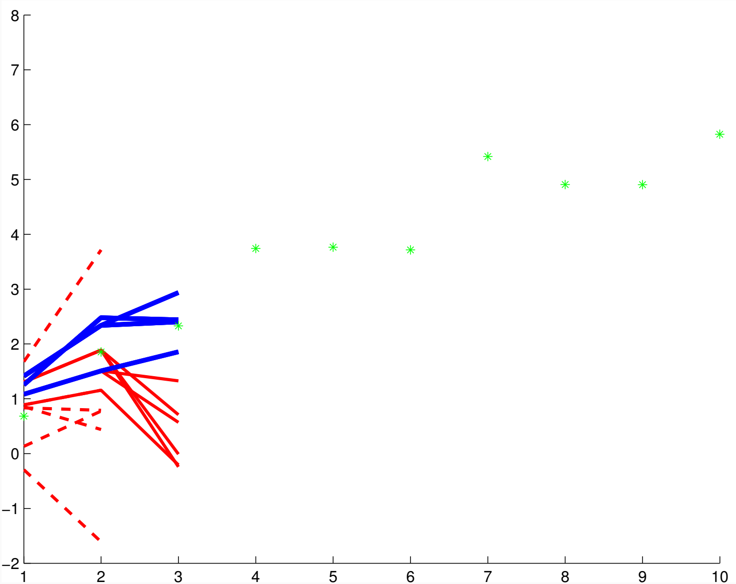
$$f(x_t|x_{t-1}) = \mathcal{N}(x_t; x_{t-1}, 1)$$
$$g(y_t|x_t) = \mathcal{N}(y_t; x_t, 1)$$

N.B. Don't do this! In the *linear Gaussian* case the Kalman Filtering recursions give exact solutions Kalman (1960).

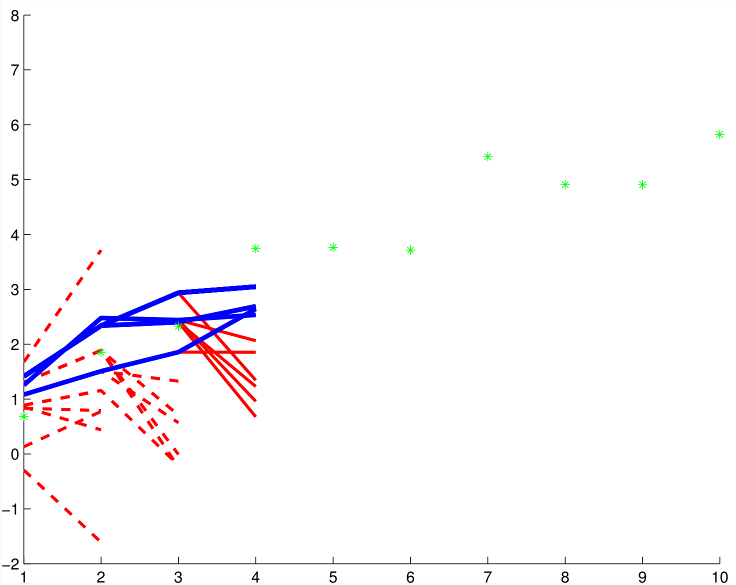
Iteration 2



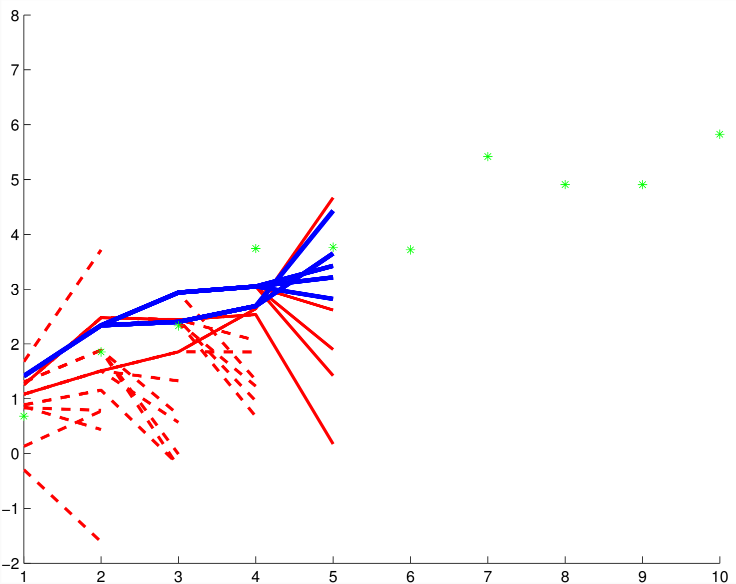
Iteration 3



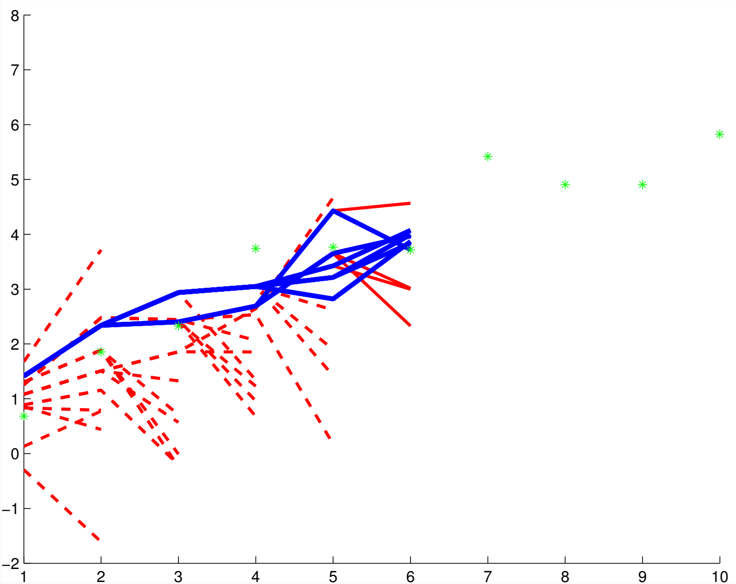
Iteration 4



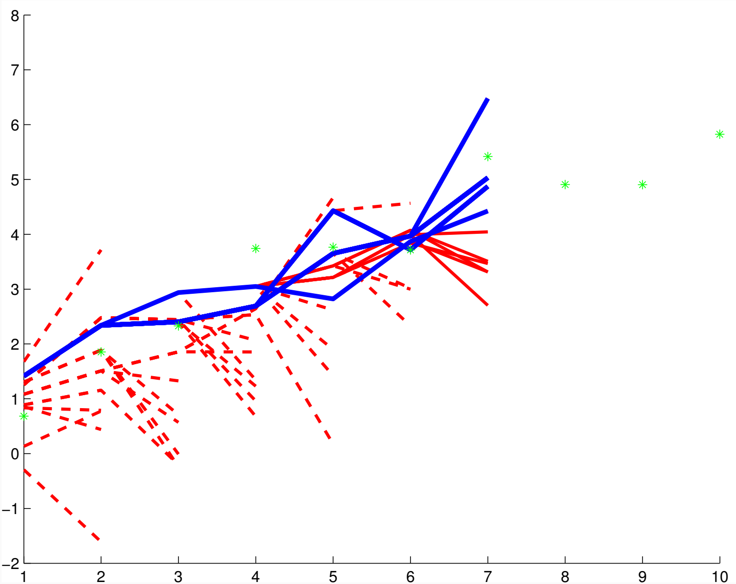
Iteration 5



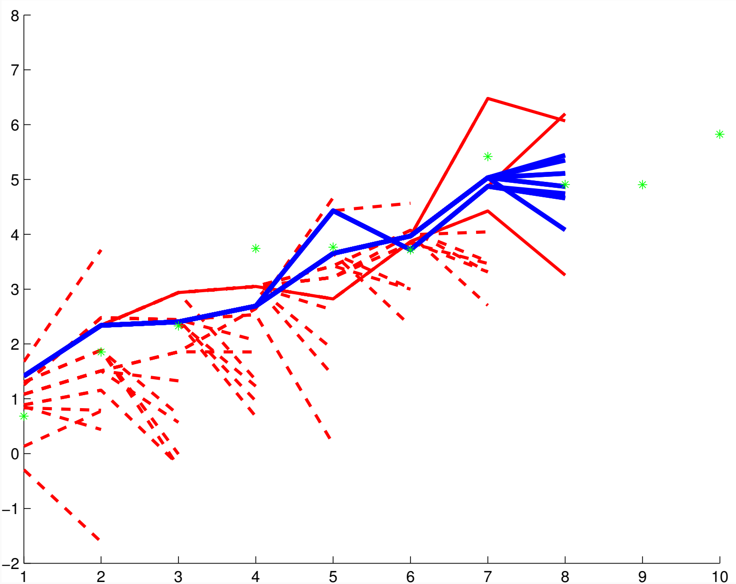
Iteration 6



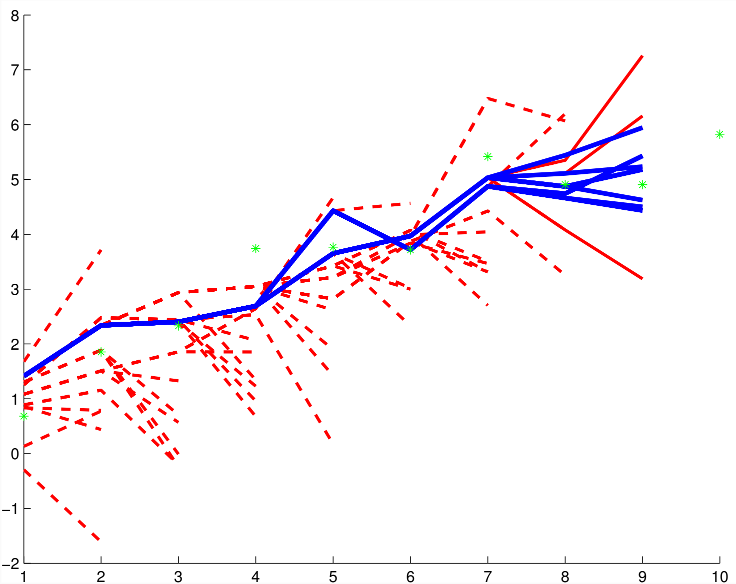
Iteration 7



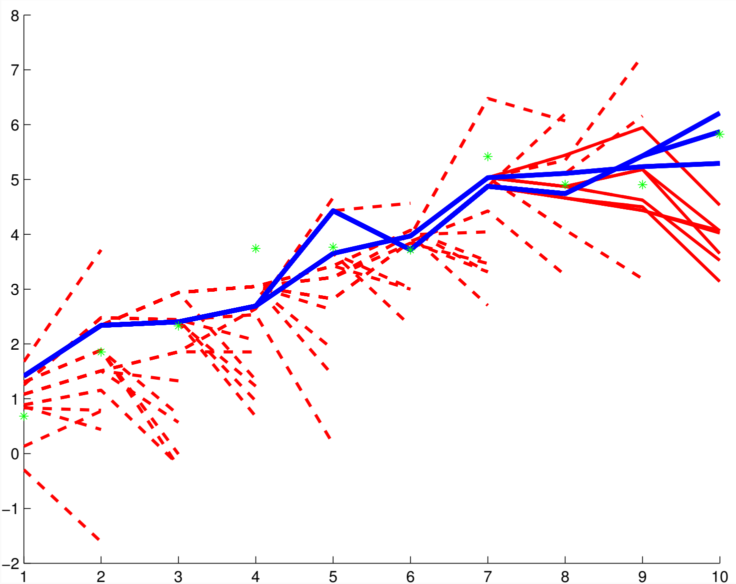
Iteration 8



Iteration 9



Iteration 10



A Modern Perspective

- The bootstrap filter (and many other algorithms) are a principled combination of *importance sampling* and *resampling*;
- this perspective admits very considerable generalization;
- and the rigorous establishment of good convergence properties (see later).

See, e.g., Doucet et al. (2000); Doucet and Johansen (2011); Chopin and Papaspiliopoulos (2020) or for a more abstract treatment Del Moral (2004).

Importance Sampling

- Simple identity: provided $\gamma \ll \mu$:

$$Z = \int \gamma(x)dx = \int \frac{\gamma(x)}{\mu(x)}\mu(x)dx$$

- So, if $X_1, X_2, \ldots \overset{\text{iid}}{\sim} \mu$, then:

$$\textbf{unbiasedness} \qquad \forall N : \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\frac{\gamma(X_i)}{\mu(X_i)}\right] = Z$$

$$\textbf{slln} \qquad \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\frac{\gamma(X_i)}{\mu(X_i)}\varphi(X_i) \overset{\text{a.s.}}{\to} \gamma(\varphi)$$

$$\textbf{clt} \qquad \lim_{N\to\infty}\sqrt{N}\left[\frac{1}{N}\sum_{i=1}^{N}\frac{\gamma(X_i)}{\mu(X_i)}\varphi(X_i) - \gamma(\varphi)\right] \overset{\text{d}}{\to} W$$

where $W \sim \mathcal{N}\left(0, \text{Var}\left[\frac{\gamma(X_1)}{\mu(X_1)}\varphi(X_1)\right]\right)$.

Sequential Importance Sampling

- Write

$$\gamma(x_{1:n}) = \gamma(x_1) \prod_{p=2}^{n} \gamma(x_p | x_{1:p-1}),$$

- define, for $p = 1, \ldots, n$

$$\gamma_p(x_{1:p}) = \gamma_1(x_1) \prod_{q=2}^{p} \gamma(x_q | x_{1:q-1}),$$

- then

$$\underbrace{\frac{\gamma(x_{1:n})}{\mu(x_{1:n})}}_{W_n(x_{1:n})} = \underbrace{\frac{\gamma_1(x_1)}{\mu_1(x_1)}}_{w_1(x_1)} \prod_{p=2}^{n} \underbrace{\frac{\gamma_p(x_{1:p})}{\gamma_{p-1}(x_{1:p-1})\mu_p(x_p|x_{1:p-1})}}_{w_p(x_{1:p})},$$

- and we can *sequentially* approximate $Z_p = \int \gamma_p(x_{1:p}) dx_{1:p}$ and $\int \pi_p(dx_p)\varphi(x_p) = Z_p^{-1} \int \gamma_p(dx_p)\varphi(x_p)$.

Resampling

- Given a weighted empirical measure:

$$\eta^N = \sum_{i=1}^{N} W^i \delta_X^i$$

- The unwegihted empirical measure

$$\tilde{\eta}^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{Y^i}$$

will satisfy, for all η^N-integrable φ,

$$\mathbb{E}\left[\int \tilde{\eta}^N(dy)\varphi(y) \mid (W^i, X^i)_{i=1}^N\right] = \int \eta^N(dx)\varphi(x), \text{ if}$$

- For $i = 1 : N : Y_i \in \{X_1, \dots, X_n\}$ w.p. 1.
- For $i = 1 : N$: $\mathbb{E}\left[N_i \mid (W^i, X^i)_{i=1}^N\right] = NW_i$, with $N_i := \#\{j : Y_i = X_j\}$.

- Simplest option: $Y^i \overset{\text{iid}}{\sim} \eta^N$.

Sequential Importance Resampling (SIR)

Given a sequence $\gamma_1(x_1), \gamma_2(x_{1:2}), \ldots$:

Initialisation, $n = 1$**:**

- Sample $X_1^1, \ldots, X_1^N \overset{\text{iid}}{\sim} \mu_1$
- Compute

$$W_1^i = \frac{\gamma_1(X_1^i)}{\mu_1^i(X_1^i)}$$

- Obtain $\widehat{Z}_1^N = \frac{1}{N}\sum_{i=1}^N W_1^i \qquad \widehat{\pi}_1^N = \frac{\sum_{i=1}^N W_1^i \delta_{X_1^i}}{\sum_{j=1}^N W_1^j}$

[This is *just* (self-normalized) importance sampling.]

Iteration, $n \leftarrow n+1$**:**

- Resample: sample $(X^1_{n,1:n-1}, \ldots, X^N_{n,1:n-1}) \overset{\text{iid}}{\sim} \sum_{i=1}^N \delta_{X^i_{n-1}}$
- Sample $X^i_{n,n} \sim q_n(\cdot | X^i_{n,1:n-1})$
- Compute

$$W^i_n = \frac{\gamma_n(X^i_{n,1:n})}{\gamma_{n-1}(X^i_{n,1:n-1}) \cdot q_n(X^i_{n,n} | X^i_{n,1:n-1})}.$$

- Obtain

$$\widehat{Z}^N_n = \widehat{Z}^N_{n-1} \cdot \frac{1}{N} \sum_{i=1}^N W^i_n \qquad \widehat{\pi}^N_n = \frac{\sum_{i=1}^N W^i_n \delta_{X^i_n}}{\sum_{j=1}^N W^j_n}.$$

Interpreting the bootstap particle filter as SIR

- Take $\gamma_n(x_{1:n}) = p(x_{1:n}, y_{1:n})$; and $q_n(x_n | x_{1:n-}1) = f(x_n | x_{n-1})$.
- recover $\frac{\gamma_n(X^i_{n,1:n})}{\gamma_{n-1}(X^i_{n,1:n-1}) \cdot q_n(X^i_{n,n} | X^i_{n,1:n-1})} = g(y_n | x_n)$.

Methodology: Computational Statistics

Better Filtering

Better proposal distributions E.g. take

$$q_n(x_n|x_{n-1}) = p(x_n|x_{n-1}, y_n) \propto f(x_n|x_{n-1})g(y_n|x_n)$$

Doucet et al. (2000).

Auxiliary/lookahead methods Auxiliary Particle Filters Pitt and Shephard (1999);Carpenter et al. (1999); Johansen and Doucet (2008) / Lookahead methods Lin et al. (2013)

Better resampling schemes see later.

Incorporating MCMC Gilks and Berzuini (2001) roughening revisited; Berzuini et al. (1997) using MCMC rather than simple sampling.

Beyond Filtering

Smoothing See Briers et al. (2010)

- **Brute force**
- **Fixed-lag**
- **Forward-filtering backward-smoothing**
- **Backward information filters**
- **Iterated Methods** iAPF Guarniero et al. (2017) / controlled SMC Heng et al. (2020)

Parameter Estimation See Kantas et al. (2015). Paticularly noteworthy: Particle MCMC Andrieu et al. (2010)

Beyond HMMs/FK models

Sampling from an arbitary sequence, or a single distribution, by constructing an artificial flow: Chopin (2002); Del Moral et al. (2006). (Closely connected to Annealed Importance Sampling; Neal (2001)).

In particular, the target densities, $\tilde{\gamma}_t$, proposals K_t and associated importance weights:

$$\begin{aligned}\tilde{\gamma}_t(x_{1:t}) =& \gamma_t(x_t)\prod_{s=1}^{t-1} L_s(x_{s+1}, x_s)\\ W_t(x_{1:t}) =& \frac{\tilde{\gamma}_t(x_{1:t})}{\tilde{\gamma}_{t-1}(x_{1:t-1})K_t(x_{t-1}, x_t)}\\ =& \frac{\gamma_t(x_t)L_{t-1}(x_t, x_{t-1})}{\gamma_{t-1}(x_{t-1}), K_t(x_{t-1}, x_t)}\end{aligned}$$

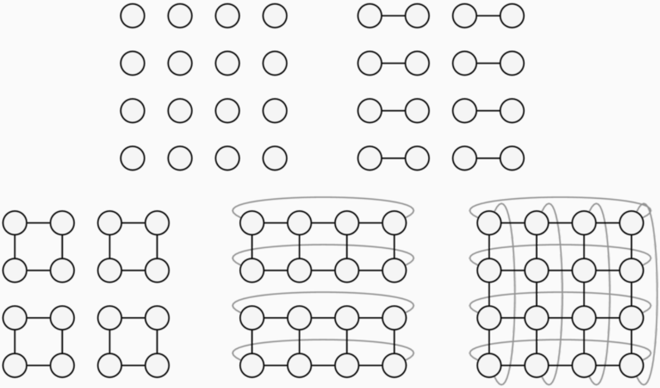
The particle-MCMC method of Andrieu et al. (2010) generalises to these contexts.

Parallel/Distributed Methods: Resampling

- Resampling is fundamentally synchronous.
- Various schemes have been proposed to circumvent this:
  - "Metropolis Resampling" Murray et al. (2016)
  - "Forest Resampling" Lee and Whiteley (2016)
  - Reduced-interaction resampling Whiteley et al. (2016)
  - Island particle methods Vergé et al. (2015)

Parallel/Distributed Methods: Divide-and-Conquer SMC

Rough idea Lindsten et al. (2017)
If "$\eta_1^N \approx \eta_1$" and "$\eta_2^N \approx \eta_2$" then "$\eta_1^N \times \eta_2^N \approx \eta_1 \times \eta_2$" + SMC.



Theory: Applied Probability & Theoretical Statistics

Feynman-Kac Del Moral (2004, 2013) and references therein for a *tour de force* treatment; see also Chopin and Papaspiliopoulos (2020) for a gentler introduction, or Del Moral, Piere and Doucet, Arnaud (2014) for a presentation on discrete spaces.

triangular arrays some general results Douc and Moulines (2008)

Properties

stability non-compact settings Whiteley (2013)

unbiasedness

lln Del Moral (1996); Crisan and Doucet (2002)

clt Del Moral and Guionnet (1999); Chopin (2004); Künsch (2005)

adaptivity resampling Del Moral et al. (2012); algorithms Beskos et al. (2016)

Feynman-Kac Flows

Feynman-Kac Model Ingredients & Recipe

State space(s) $(E_t)t \geq 0$ (often $E_t \equiv t$)

Markov chain with kernels $K_t(x_{t-1}, dx_t)$ from E_{t-1} to E_t

Potentials $(G_t : E_t \to [0, \infty))_{t \geq 0}$

Markov chain $\mathbb{P}(X_{1:t} \in A) = \int_A \eta_1(dx_1) \prod_{p=2}^t K_p(x_{p-1}, dx_p)$

FK distributions

$$\mathbb{Q}_t(X_{1:t} \in A) = \frac{\int_A \eta_1(dx_1) \prod_{p=2}^t [K_p(x_{p-1}, dx_p) G_{p-1}(x_{p-1})]}{\int_{E_1 \times \cdots \times E_p} \eta_1(dx_1) \prod_{p=2}^t [K_p(x_{p-1}, dx_p) G_{p-1}(x_{p-1})]}$$

Marginals

$$\eta_t(dx_t) = \mathbb{Q}_t(X_t \in dx_t) = \frac{\int \eta_{t-1}(dx_{t-1}) G_{t-1}(x_{t-1}) M_t(x_{t-1}, dx_t)}{\int \eta_{t-1}(dx_{t-1}) G_{t-1}(x_{t-1})}$$

A Feynman-Kac Representation of Filtering

Ingredients

Transitions $M_t(x_{t-1}, dx_t) = f(x_t|x_{t-1})dx_t$ (and $\eta_1 = \mu_1$)

Potentials $G_t(x_t) = g(y_t|x_t)$ ($y_1, \ldots$ fixed)

Recursion

$$\eta_t(dx_t) \propto \int \mu(x_1) \prod_{p=2}^{t} \left[f(x_p|x_{p-1}) g(y_{p-1}|x_{p-1}) \right] dx_{1:t}$$

$$= \int \eta_{t-1}(dx_{t-1}) g(y_t|x_t) f(x_t|x_{t-1}) dx_t$$

prediction $\eta_t(dx_t) = \hat{\eta}_{t-1}(dx_{t-1}) M_t(x_{t-1}, dx_t)$

update $\hat{\eta}_t(dx_t) := G(x_t)\eta_t(dx_t) / \int G(x_t')\eta_t(dx_t')$

A Mean-field Particle Approximation

Initialization
Sample $\{X_1^i\}_{i=1}^N \sim \eta_1$; set $\eta_1^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_1^i}$.

Recursion, $t = 2, \ldots$

update Sample N times from

$$\tilde{\eta}_{t-1}^N(dx_{t-1}) = \frac{G_{t-1}(x_{t-1})\hat{(}dx_{t-1})}{\int \hat{\eta}_{t-1}^N(dx'_{t-1})G_{t-1}(x'_{t-1})} = \frac{\sum_{i=1} G_{t-1}(X_{t-1}^i)\delta_{X_{t-1}^i}}{\sum_{jj=1} G_{t-1}(X_{t-1}^)}$$

to obtain $\{\tilde{X}_{t-1}^i\}_{i=1}^N$ and $\hat{\eta}_{t-1}^N = \frac{1}{N}\sum_{i=1}^N \delta_{\tilde{X}_{t-1}^i}$.

prediction Sample $\{X_t^i \sim M_t(\tilde{X}_{t-1}^i, \cdot)\}_{i=1}^N$ — a stratified sample from $\int \hat{\eta}^N(dx_{t-1})M_t(x_{t-1}, \cdot)$.

Local Error Decomposition

$$
\begin{array}{ccccccccc}
\eta_1 & \rightarrow & \eta_2 = \Phi_2(\eta_1) & \rightarrow & \eta_3 = \Phi_{1:3}(\eta_1) & \rightarrow & \ldots & \rightarrow \Phi_{1:n}(\eta_1) \\
\Downarrow & & & & & & & \\
\eta_1^N & \rightarrow & \Phi_2(\eta_1^N) & \rightarrow & \Phi_{1:3}(\eta_1^N) & \rightarrow & \ldots & \rightarrow \Phi_{1:n}(\eta_1^N) \\
 & & \Downarrow & & & & & \\
 & & \eta_2^N & \rightarrow & \Phi_3(\eta_2^N) & \rightarrow & \ldots & \rightarrow \Phi_{2:n}(\eta_2^N) \\
 & & & & \Downarrow & & & \\
 & & & & \eta_3^N & \rightarrow & \ldots & \rightarrow \Phi_{3:n}(\eta_3^N) \\
 & & & & & & \vdots & \\
 & & & & & & \Downarrow & \\
 & & & & & & \eta_{n-1}^N & \rightarrow \Phi_n(\eta_{n-1}^N) \\
 & & & & & & & \Downarrow \\
 & & & & & & & \eta_n^N
\end{array}
$$

n

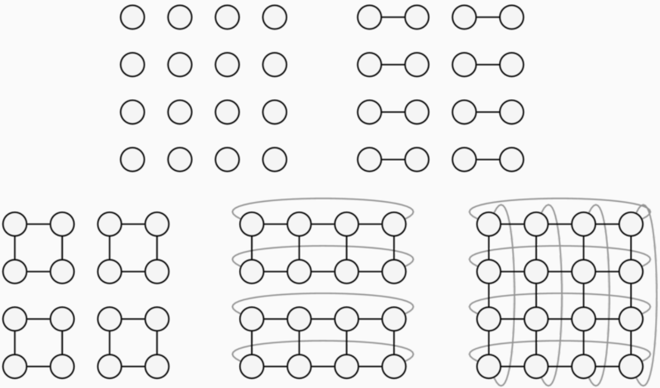
One Recent Example: Justifying MCMC particle filters

- Introduced in Berzuini et al. (1997);
- strong law of large numbers and clt provided by Finke et al. (2020); conditions required for clt are prohibitively strong;
- inspiring a different approach. RSN: Caprio and Johansen (2023)

Another Recent Example: Justifying D&C-SMC i

Recall Idea

If "$\eta_1^N \approx \eta_1$" and "$\eta_2^N \approx \eta_2$" then "$\eta_1^N \times \eta_2^N \approx \eta_1 \times \eta_2$" + SMC.



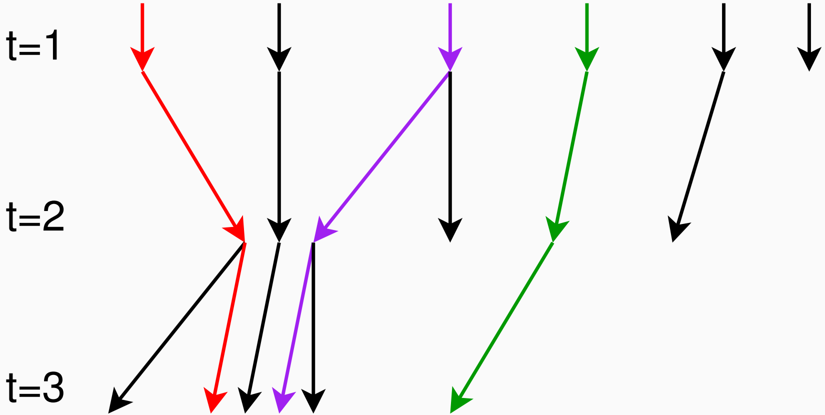
Another Recent Example: Justifying D&C-SMC ii

- Initial idea Lindsten et al. (2017) with rudimentary justification.
- Comprehensive justification Kuntz et al. (2023). . .
- inspiring a study of product measures in general Kuntz et al. (2022); with an interesting connection to multi-sample U-statistics!
- In turn provides a theoretical framework for other algorithms, such as "Divide and Conquer Fusion" Chan et al. (2023).
- Finds application in smoothing Corneflos et al. (2022), high-dimensional filtering Crucinio and Johansen (2023)

Resampling

- Early methodological contributions: residual resampling Liu and Chen (1998), systematic resampling Carpenter et al. (1999), stratified resamping Kitagawa (1996), tree-based branching approximations Crisan and Lyons (2002)
- A basic theoretical assessment: Douc et al. (2005)
- A more comprehensive theory: Gerber et al. (2019)

Ancestral Trees



$a_2^1 = 1$ $a_2^4 = 3$ $a_1^1 = 1$ $a_1^4 = 3$

$b_{3,1:3}^2 = (1, 1, 2)$ $b_{3,1:3}^4 = (3, 3, 4)$ $b_{3,1:3}^6 = (4, 5, 6)$

A Connection to Population Biology...

- Del Moral et al. (2009): asymptotic genealogy of neutral models
- Jacob et al. (2015): stochastic finite sample bounds
- Koskela et al. (2020); Brown et al. (2021, 2023): asymptotic genealogy of *non-neutral* models

Some Applications: Applied Statistics, Econometrics, Engineering...

Some grains of sand from a desert.

Target tracking defence; air traffic control; navigation

Navigation Mars rover Ng et al. (2005)

SLAM Simultaneous Localisation and Mapping Bailey and Durrant-Whyte (2006)

Econometrics Numerous applications see Lopes and Tsay (2011) and references therein.

Battery Life Monitoring Liu et al. (2011)

Neuroscience Neuroimaging Aston and Johansen (2015)

Summary

- From small and unassuming papers, “great” things can develop!
- By popularizing an approach to Monte Carlo simulation in state-space models, this paper was instrumental in precipitating a lot of diverse research.
- An “engineering paper” is one the of things behind development of methodology and theory in inferential and computational statistics, strands of research in applied probability and it underpins a vast range of more applied research.
- This talk has only scratched the surface.

References i

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

J. A. D. Aston and A. M. Johansen. Bayesian inference on the brain: Bayesian solutions to selected problems in neuroimaging. In S. K. Upadhyay, U. Singh, D. K. Dey, and A. Loganathan, editors, *Current Trends in Bayesian Methodology*, chapter 1, pages 1–36. CRC Press, 2015. doi: 10.1201/b18502-2.

T. Bailey and H. Durrant-Whyte. Simultaneous localization and mapping: Part II state of the art. *IEEE Robotics and Automation Magazine*, 13(3):108–117, September 2006.

C. Berzuini, N. G. Best, W. R. Gilks, and C. Larizza. Dynamic conditional independence models and Markov Chain Monte Carlo. *Journal of the American Statistical Association*, 92(440):1403–1412, December 1997. URL http://links.jstor.org/sici?sici=0162-1459%28199712%2992%3A440%3C1403%3ADCIMAM%3E2.0.CO%3B2-U.

A. Beskos, A. Jasra, N. Kantas, and A. Thiery. On the convergence of adaptive sequential monte carlo methods. *The Annals of Applied Probability*, 26(2):1111–1146, 2016.

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.

S. Brown, P. A. Jenkins, A. M. Johansen, and J. Koskela. Simple conditions for convergence of sequential Monte Carlo genealogies with applications. *Electronic Journal of Probability*, 26:1–22, 2021. ISSN 1083-6489. doi: 10.1214/20-EJP561.

S. Brown, P. Jenkins, A. M. Johansen, and J. Koskela. Weak convergence of non-neutral genealogies to Kingman's coalescent. *Stochastic Processes and their Applications*, 162C:76–105, 2023. doi: 10.1016/j.spa.2023.04.016.

R. Caprio and A. M. Johansen. A calculus for Markov chain Monte Carlo: studying approximations in algorithms. In preparation, 2023.

J. Carpenter, P. Clifford, and P. Fearnhead. An improved particle filter for non-linear problems. *Proceedings on Radar, Sonar and Navigation*, 146(1):2–7, 1999.

References ii

R. Chan, M. Pollock, A. M. Johansen, and G. O. Roberts. Divide-and-conquer Monte Carlo fusion. *Journal of Machine Learning Research*, 24(193):1–82, June 2023. URL http://jmlr.org/papers/v24/21-1274.html.

N. Chopin. A sequential particle filter method for static models. *Biometrika*, 89(3):539–551, 2002.

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.

A. Corneflos, N. Chopin, and S. Särkkä. De-Sequentialized Monte Carlo: a parallel-in-time particle smoother. *Journal of Machine Learning Research*, 23(283):1–39, 2022. URL http://jmlr.org/papers/v23/22-0140.html.

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.

D. Crisan and T. Lyons. Minimal entropy approximations and optimal algorithms. *Mont Carlo Methods and Applications*, 8(4):343–356, 2002. doi: doi:10.1515/mcma.2002.8.4.343.

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.

P. Del Moral. Nonlinear filtering: interacting particle solution. *Markov Processes and Related Fields*, 2(4): 555–580, 1996.

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 A. Guionnet. Central limit theorem for non linear filtering and interacting particle systems. *Annals of Applied Probability*, 9(2):275–297, 1999.

P. Del Moral, A. Doucet, and A. Jasra. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society B*, 63(3):411–436, 2006. doi: 10.1111/j.1467-9868.2006.00553.x.

References iii

P. Del Moral, L. Miclo, F. Patras, and S. Rubenthaler. The convergence to equilibrium of neutral genetic models. *Stochastic Analysis and Applications*, 28(1):123–143, 2009. doi: 10.1080/07362990903415833.

P. Del Moral, A. Doucet, and A. Jasra. On adaptive resampling procedures for sequential Monte Carlo methods. *Bernoulli*, 18(1):252–278, 2012. URL http://hal.inria.fr/inria-00332436/PDF/RR-6700.pdf.

Del Moral, Piere and Doucet, Arnaud. Particle methods: An introduction with applications. *ESAIM: Proc.*, 44: 1–46, 2014. doi: 10.1051/proc/201444001. URL https://doi.org/10.1051/proc/201444001.

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.

R. Douc, O. Cappé, and E. Moulines. Comparison of resampling schemes for particle filters. In *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis*, volume I, pages 64–69. IEEE, 2005.

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

A. Finke, A. Doucet, and A. M. Johansen. Limit theorems for sequential MCMC methods. *Advances in Applied Probability*, 52(2):377–403, 2020. doi: 10.1017/apr.2020.9.

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

W. R. Gilks and C. Berzuini. Following a moving target – Monte Carlo inference for dynamic Bayesian models. *Journal of the Royal Statistical Society B*, 63:127–146, 2001.

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.

References iv

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

P. E. Jacob, L. Murray, and S. Rubenthaler. Path storage in the particle filter. *Statistics and Computing*, 25(2): 487–496, 2015.

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.

R. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82:35–42, 1960.

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.

G. Kitagawa. Monte Carlo filter and smoother for Non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5:1–25, 1996.

J. Koskela, P. Jenkins, A. M. Johansen, and D. Spanò. Asymptotic genealogies of interacting particle systems with an application to sequential Monte Carlo. *Annals of Statistics*, 48(1):560–583, 2020. doi: 10.1214/19-AOS1823.

H. R. Künsch. Recursive Monte Carlo filters: Algorithms and theoretical analysis. *Annals of Statistics*, 33(5): 1983–2021, 2005.

J. Kuntz, F. R. Crucinio, and A. M. Johansen. Product-form estimators: exploiting independence to scale up Monte Carlo. *Statistics and Computing*, 32(12):1–22, 2022. doi: 10.1007/s11222-021-10069-9.

J. Kuntz, F. R. Crucinio, and A. M. Johansen. Divide-and-conquer sequential Monte Carlo: Properties and limit theorems. *Annals of Applied Probability*, 2023. In press.

A. Lee and N. Whiteley. Forest resampling for distributed sequential Monte Carlo. *Statistical Analysis and Data Mining*, 9(4):230–248, 2016.

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

References v

F. Lindsten, A. M. Johansen, C. A. Naesseth, B. Kirkpatrick, T. Schön, J. A. D. Aston, and A. Bouchard-Côté. Divide and conquer with sequential Monte Carlo samplers. *Journal of Computational and Graphical Statistics*, 26(2):445–458, 2017. doi: 10.1080/10618600.2016.1237363.

J. Liu, W. Wang, and F. Ma. A regularized auxiliary particle filtering approach for system state estimation and battery life prediction. *Smart Materials and Structures*, 20(075021):9, 2011.

J. S. Liu and R. Chen. Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93(443):1032–1044, September 1998.

H. F. Lopes and R. S. Tsay. Particle filters and bayesian inference in financial econometrics. *Journal of Forecasting*, 30(1):168–209, 2011. doi: https://doi.org/10.1002/for.1195.

L. Murray, A. Lee, and P. Jacob. Parallel resampling in the particle filter. *Journal of Computational and Graphical Statistics*, 25(3):789–805, 2016.

R. M. Neal. Annealed importance sampling. *Statistics and Computing*, 11:125–139, 2001.

B. Ng, A. Pfeffer, and R. Dearden. Continuous time particle filtering. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, August 2005.

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

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.

C. Vergé, C. Dubarry, P. Del Moral, and E. Moulines. On parallel implementation of sequential monte carlo methods: the island particle model. *Statistics and Computing*, 25(2):243–260, 2015.

N. Whiteley. Sequential monte carlo samplers: error bounds and insensitivity to initial conditions. *Stochastic Analysis and Applications*, 30(5):774–798, 2013. In press.

N. Whiteley, A. Lee, and K. Heine. On the role of interaction in sequential Monte Carlo algorithms. *Bernoulli*, 22 (1):494–429, 2016.