Controlled Sequential Monte Carlo
arXiv:1708.08396

Jeremy Heng

Department of Statistics, Harvard University
Joint work with Adrian Bishop (UTS & CSIRO), George Deligiannidis (KCL) and Arnaud Doucet (Oxford)

SMC Workshop 2017

Feynman-Kac path measure

- Consider a non-homogenous Markov chain $(X_t)_{t\in[0:T]}$ on $(\mathsf{X},\mathcal{X})$ with law

$$\mathbb{Q}(dx_{0:T}) = \mu(dx_0)\prod_{t=1}^{T} M_t(x_{t-1}, dx_t)$$

- Given positive bounded potential functions $(G_t)_{t\in[0:T]}$, define **Feynman-Kac path measure**

$$\mathbb{P}(dx_{0:T}) = G_0(x_0)\prod_{t=1}^{T} G_t(x_{t-1}, x_t)\mathbb{Q}(dx_{0:T})Z^{-1}$$

  where $Z := \mathbb{E}_{\mathbb{Q}}\left[G_0(X_0)\prod_{t=1}^{T} G_t(X_{t-1}, X_t)\right]$

- The quantities $\left\{\mu, (M_t)_{t\in[1:T]}, (G_t)_{t\in[0:T]}\right\}$ depend on the specific application
- Applications of interest: **static models** and **state space models**

Sequential Monte Carlo methods

- SMC methods simulate an **interacting particle system** of size $N \in \mathbb{N}$
- At time $t = 0$ and particle $n \in [1 : N]$
  - sample $X_0^n \sim \mu$;
  - sample ancestor index $A_0^n \sim \mathcal{R}\left(G_0(X_0^1), \ldots, G_0(X_0^N)\right)$
- For time $t \in [1 : T]$ and particle $n \in [1 : N]$
  - sample $X_t^n \sim M_t(X_{t-1}^{A_{t-1}^n}, \cdot)$;
  - sample ancestor index $A_t^n \sim \mathcal{R}\left(G_t(X_{t-1}^{A_{t-1}^1}, X_t^1), \ldots, G_t(X_{t-1}^{A_{t-1}^N}, X_t^N)\right)$

Sequential Monte Carlo methods

- Particle approximation of $\mathbb{P}$

$$\mathbb{P}^N = \frac{1}{N}\sum_{n=1}^{N}\delta_{X_{0:T}^n}$$

where $X_{0:T}^n$ is obtained by tracing ancestral lineage of particle X_T^n

- Unbiased estimator of Z

$$Z^N = \left\{\frac{1}{N}\sum_{n=1}^{N} G_0(X_0^n)\right\}\prod_{t=1}^{T}\left\{\frac{1}{N}\sum_{n=1}^{N} G_t(X_{t-1}^{A_{t-1}^n}, X_t^n)\right\}$$

- Convergence properties of $\mathbb{P}^N$ and Z^N as $N \to \infty$ are now well-understood
- However quality of approximation can be inadequate for practical choices of N
- Performance crucially depends on discrepancy between $\mathbb{P}$ and $\mathbb{Q}$

Twisted path measures

- Consider **change of measure** prescribed by positive and bounded functions $\psi = (\psi_t)_{t\in[0:T]}$
- Refer to ψ as an **admissible policy** and denote **set of all admissible policies** as Ψ
- Given a policy $\psi \in \Psi$, define ψ-**twisted path measure** of $\mathbb{Q}$ as

$$\mathbb{Q}^{\psi}(dx_{0:T}) = \mu^{\psi}(dx_0)\prod_{t=1}^{T} M_t^{\psi}(x_{t-1}, dx_t)$$

where

$$\mu^{\psi}(dx_0) := \frac{\mu(dx_0)\psi_0(x_0)}{\mu(\psi_0)}, \quad M_t^{\psi}(x_{t-1}, dx_t) := \frac{M_t(x_{t-1}, dx_t)\psi_t(x_{t-1}, x_t)}{M_t(\psi_t)(x_{t-1})},$$

for $t \in [1:T]$

Twisted path measures

- Given $\psi \in \Psi$, we have

$$\mathbb{P}(\mathrm{d}x_{0:T}) = G_0^{\psi}(x_0) \prod_{t=1}^{T} G_t^{\psi}(x_{t-1}, x_t)\, \mathbb{Q}^{\psi}(dx_{0:T}) Z^{-1}$$

where

$$G_0^{\psi}(x_0) := \frac{\mu(\psi_0) G_0(x_0) M_1(\psi_1)(x_0)}{\psi_0(x_0)},$$

$$G_t^{\psi}(x_{t-1}, x_t) := \frac{G_t(x_{t-1}, x_t) M_{t+1}(\psi_{t+1})(x_t)}{\psi_t(x_{t-1}, x_t)}, \quad t \in [1:T-1],$$

$$G_T^{\psi}(x_{T-1}, x_T) := \frac{G_T(x_{T-1}, x_T)}{\psi_T(x_{T-1}, x_T)},$$

are the **twisted potentials** associated with $\mathbb{Q}^{\psi}$

- Note $Z = \mathbb{E}_{\mathbb{Q}^{\psi}} \left[G_0^{\psi}(X_0) \prod_{t=1}^{T} G_t^{\psi}(X_{t-1}, X_t) \right]$ by construction

Twisted SMC methods

- Assume policy $\psi \in \Psi$ is such that:
  - sampling μ^ψ and $(M_t^\psi)_{t\in[1:T]}$ feasible
  - evaluating $(G_t^\psi)_{t\in[0:T]}$ tractable
- Construct ψ-**twisted SMC method** as standard SMC applied to $\left\{\mu^\psi, (M_t^\psi)_{t\in[1:T]}, (G_t^\psi)_{t\in[0:T]}\right\}$
- Particle approximation of $\mathbb{P}$ and Z

$$\mathbb{P}^{\psi,N} = \frac{1}{N}\sum_{n=1}^{N} \delta_{X_{0:T}^n}, \quad Z^{\psi,N} = \left\{\frac{1}{N}\sum_{n=1}^{N} G_0^\psi(X_0^n)\right\} \prod_{t=1}^{T} \left\{\frac{1}{N}\sum_{n=1}^{N} G_t^\psi(X_{t-1}^{A_{t-1}^n}, X_t^n)\right\}$$

Optimal policies

- A policy with constant functions recover standard SMC method
- Consider an iterative scheme to refine policies
- Given current policy $\psi \in \Psi$, twisting $\mathbb{Q}^{\psi}$ further with policy $\phi \in \Psi$ results in a twisted path measure $(\mathbb{Q}^{\psi})^{\phi}$
- Note that $(\mathbb{Q}^{\psi})^{\phi} = \mathbb{Q}^{\psi \cdot \phi}$ where $\psi \cdot \phi = (\psi_t \cdot \phi_t)_{t \in [0:T]}$
- Choice of ϕ is guided by the following optimality result

Optimal policies

Proposition

For any $\psi \in \Psi$, under the policy $\phi^* = (\phi_t^*)_{t \in [0:T]}$ defined recursively as

$$\phi_T^*(x_{T-1}, x_T) = G_T^\psi(x_{T-1}, x_T),$$
$$\phi_t^*(x_{t-1}, x_t) = G_t^\psi(x_{t-1}, x_t) M_{t+1}^\psi(\phi_{t+1}^*)(x_t), \quad t \in [T-1 : 1],$$
$$\phi_0^*(x_0) = G_0^\psi(x_0) M_1^\psi(\phi_1^*)(x_0),$$

the refined policy $\psi^* := \psi \cdot \phi^*$ satisfies:

(i) $\mathbb{P} = \mathbb{Q}^{\psi^*}$;

(ii) $Z^{\psi^*, N} = Z$ *almost surely for any* $N \in \mathbb{N}$.

Optimal policies

- Refer to ϕ^* as the optimal policy w.r.t. $\mathbb{Q}^{\psi}$
- The refined policy $\psi^* = \psi \cdot \phi^*$ is the optimal policy w.r.t. $\mathbb{Q}$
- ψ^*-twisted potentials

$$G_0^{\psi^*}(x_0) = Z, \quad G_t^{\psi^*}(x_{t-1}, x_t) = 1, \quad t \in [1:T]$$

- Under ψ^*-twisted SMC method

$$Z_t^{\psi^*,N} = \left\{ \frac{1}{N} \sum_{n=1}^{N} G_0^{\psi^*}(X_0^n) \right\} \prod_{k=1}^{t} \left\{ \frac{1}{N} \sum_{n=1}^{N} G_k^{\psi^*}(X_{k-1}^{A_{t-1}^n}, X_k^n) \right\} = Z$$

for all $t \in [0:T]$

Optimal policies

- The connection to Kullback-Leibler optimal control is given by

Proposition

The functions $V_t^* := -\log \phi_t^*, t \in [0:T]$ are the optimal value functions of the KL control problem

$$\inf_{\phi \in \Phi} \mathrm{KL}\left((\mathbb{Q}^{\psi})^{\phi} | \mathbb{P}\right)$$

where $\Phi := \{\phi \in \Psi : \mathrm{KL}((\mathbb{Q}^{\psi})^{\phi}|\mathbb{P}) < \infty\}$.

- The following is a characterization of ϕ^* in a specific setting

Proposition

For any policy $\psi \in \Psi$ such that the corresponding twisted potentials $(G_t^{\psi})_{t\in[0:T]}$ and transition densities of $(M_t^{\psi})_{t\in[1:T]}$ are log-concave on their domain of definition, then the optimal policy $\phi^* = (\phi_t^*)_{t\in[0:T]}$ w.r.t. $\mathbb{Q}^{\psi}$ is a sequence of log-concave functions.

Dynamic programming recursions

- Simplify notation by defining the **Bellman operators** $(Q_t^{\psi})_{t\in[0:T-1]}$
- Rewrite the backward recursion defining $\phi^* = (\phi_t^*)_{t\in[0:T]}$ as

$$\phi_T^* = G_T^{\psi},$$
$$\phi_t^* = Q_t^{\psi}\phi_{t+1}^*, \quad t \in [T-1:0]$$

where

$$Q_t^{\psi}(\varphi)(x,y) = G_t^{\psi}(x,y)M_{t+1}^{\psi}(\varphi)(y)$$

- It will be convenient to view $Q_t^{\psi} : L^2(\nu_{t+1}^{\psi}) \to L^2(\nu_t^{\psi})$ where

$$\nu_0^{\psi} := \mu^{\psi}, \quad \nu_t^{\psi}(dx,dy) := \eta_{t-1}^{\psi}(dx)M_t^{\psi}(x,dy)$$

- Need to approximate this recursion in practice

Approximate projections

- Given probability measure ν and function class $\mathsf{F} \subset L^2(\nu)$,
- Define (logarithmic) projection of f onto F as

$$P^{\nu} f = \exp\left(- \arg\min_{\varphi \in \mathsf{F}} \| \varphi - (-\log f) \|^2_{L^2(\nu)} \right), \text{ for } -\log f \in L^2(\nu)$$

- Since $V_t^* = -\log \phi_t^*$ this corresponds to learning associated value functions (more stable numerically)
- A practical implementation replaces ν with a Monte Carlo approximation ν^N
- Define approximate (F, ν)-projection as

$$P^{\nu,N} f = \exp\left(- \arg\min_{\varphi \in \mathsf{F}} \| \varphi - (-\log f) \|^2_{L^2(\nu^N)} \right)$$

Approximate dynamic programming

- To use output of ψ-twisted SMC to learn optimal ϕ^*
- Define

$$\nu_0^{\psi,N} = \frac{1}{N}\sum_{n=1}^{N} \delta_{X_0^n}, \quad \nu_t^{\psi,N} = \frac{1}{N}\sum_{n=1}^{N} \delta_{\left(X_{t-1}^{A_{t-1}^n}, X_t^n\right)}, \quad t \in [1:T],$$

  which are consistent approximations of $(\nu_t^{\psi})_{t\in[0:T]}$

- Given function class $\mathsf{F}_t \subset L^2(\nu_t^{\psi})$, denote approximate $(\mathsf{F}_t, \nu_t^{\psi})$-projection by $P_t^{\psi,N}$
- Approximate backward recursion defining $\phi^* = (\phi_t^*)_{t\in[0:T]}$ by

$$\hat{\phi}_T = P_T^{\psi,N} G_T^{\psi},$$
$$\hat{\phi}_t = P_t^{\psi,N} Q_t^{\psi} \hat{\phi}_{t+1}, \quad t \in [T-1:0]$$

- This is the **approximate dynamic programming** (ADP) algorithm for finite horizon control problems (Bertsekas and Tsitsiklis, 1996)

Policy refinement

- Construct iterative algorithm: **Controlled SMC**
  - Initialization: set $\psi^{(0)}$ as constant one functions
  - For iterations $i \in [0 : I - 1]$:
    - run $\psi^{(i)}$-twisted SMC;
    - perform ADP with SMC output to obtain policy $\hat{\phi}^{(i+1)}$;
    - construct refined policy $\psi^{(i+1)} = \psi^{(i)} \cdot \hat{\phi}^{(i+1)}$.
  - At iteration $i = I$: run $\psi^{(I)}$-twisted SMC

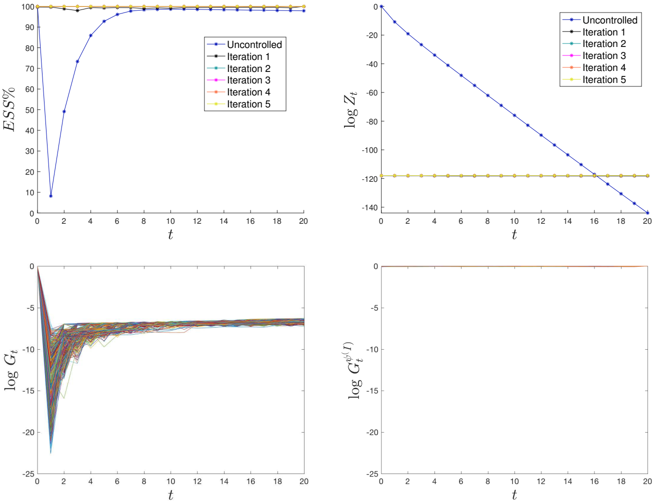


Figure: Illustration on logistic regression example.

Approximate dynamic programming

- We obtain error bounds like

$$\mathbb{E}^{\psi,N}\|\hat{\phi}_t - \phi_t^*\|_{L^2(\nu_t^\psi)} \leq \sum_{s=t}^{T} C_{t-1,s-1}^{\psi} e_s^{\psi,N}, \quad t \in [0:T]$$

  where $C_{t,s}^{\psi}$ are **stability constants** of Bellman operators and $e_t^{\psi,N}$ are **errors** of approximate projections
- As $N \to \infty$, one expects $\hat{\phi}$ to converge to $\tilde{\phi} = (\tilde{\phi}_t)_{t\in[0:T]}$, defined by the **idealized ADP** algorithm

$$\tilde{\phi}_T = P_T^{\psi} G_T^{\psi},$$
$$\hat{\phi}_t = P_t^{\psi} Q_t^{\psi} \tilde{\phi}_{t+1}, \quad t \in [T-1:0],$$

  where P_t^{ψ} is the exact $(\mathsf{F}_t, \nu_t^{\psi})$-projection
- We establish a **LLN** and **CLT** in the case where $(\mathsf{F}_t)_{t\in[0:T]}$ are given by a linear basis functions

Policy refinement

- Residuals of logarithmic projections in ADP

$$\varepsilon_t^{\psi} := \log \hat{\phi}_t - \Big(\log G_t^{\psi} - \log M_{t+1}^{\psi}(\hat{\phi}_{t+1})\Big)$$

- Related to twisted potentials of refined policy $\psi \cdot \hat{\phi}$ via

$$\log G_t^{\psi\cdot\hat{\phi}} = -\varepsilon_t^{\psi}$$

- If we twist $\mathbb{Q}^{\psi\cdot\hat{\phi}}$ further by a policy $\hat{\zeta} \in \Psi$, logarithmic projections in ADP are

$$-\log \hat{\zeta}_t := \arg\min_{\varphi \in \mathsf{F}_t} \|\varphi - (\varepsilon_t^{\psi} - \log M_{t+1}^{\psi\cdot\hat{\phi}}(\hat{\zeta}_{t+1}))\|_{L^2(\nu_t^{\psi\cdot\hat{\phi},N})}$$

where $(\nu_t^{\psi\cdot\hat{\phi},N})_{t\in[0:T]}$ are defined using output of $(\psi \cdot \hat{\phi})$-twisted SMC

Policy refinement

- Beneficial to have an iterative scheme to construct more refined policies
- Allows repeated least squares **fitting of residuals** – in the spirit of L^2-**boosting** methods
- $\mathsf{F}_t = \left\{\varphi(x_t) = x_t^T A_t x_t + x_t^T b_t + c_t : (A_t, b_t, c_t) \in \mathbb{S}_d \times \mathbb{R}^d \times \mathbb{R}\right\}$

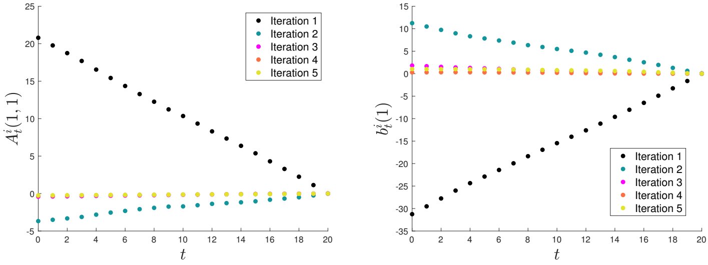


Figure: Coefficients estimated at each iteration of controlled SMC.

Iterated ADP

- Want to understand the behaviour of policy $\psi^{(I)}$ as $I \to \infty$
- Equipped Ψ with a metric ρ
- Write iterating ADP as **iterated random function** $F_U^N(\psi) = \psi \cdot \hat{\phi}$, where $\hat{\phi}$ is ADP approximation with N particles
- Iterating F^N defines a Markov chain $(\psi^{(I)})_{I \in \mathbb{N}}$ on Ψ
- Under regularity conditions, it converges to a unique invariant distribution π
- Write iterating ADP with exact projections as $F(\psi) = \psi \cdot \tilde{\phi}$, where $\tilde{\phi}$ is idealized ADP approximation
- If we assume additionally that

  $$\rho(F_U^N(\psi), F(\psi) \leq O_P(N^{-1/2})$$

  for all $\psi \in \Psi$ then

  $$\mathbb{E}_\pi \left[\rho(\psi, \varphi^*)\right] \leq O(N^{-1/2})$$

  where φ^* is a fixed point of F

Iterated ADP

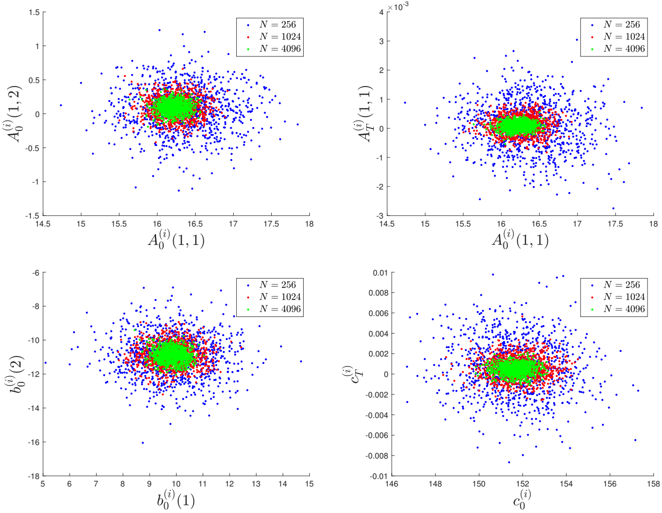


Figure: Illustrating invariant distribution of coefficients.

Log-Gaussian Cox point process

- Example from Møller et al. (1998)
- Dataset: 126 Scots pine saplings in a natural forest in Finland

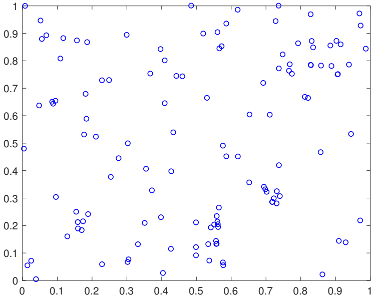


Figure: Locations of 126 Scots pine saplings in square plot of $10 \times 10\, m^2$.

Log-Gaussian Cox point process

- Discretize into a 30×30 regular grid, so $d = 900$ here
- Posterior distribution

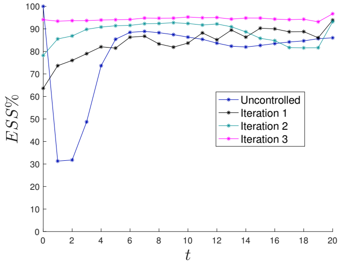
$$\eta(dx) = \mathcal{N}(x; \mu_0, \Sigma_0) \prod_{m \in [1:30]^2} \exp\left(x_m y_m - a \exp(x_m)\right) Z^{-1}$$

- Geometric path: $\eta_t(dx) = \mathcal{N}(x; \mu_0, \Sigma_0) \ell(x, y)^{\lambda_t} Z_t^{-1}$, $0 = \lambda_0 < \cdots < \lambda_T = 1$
- Set $\mu = \mathcal{N}(\mu_0, \Sigma_0)$ and M_t as unadjusted Langevin algorithm (ULA) targeting η_t
- Function classes

$$\begin{aligned} \mathsf{F}_t = \Big\{ &\varphi(x_{t-1}, x_t) = x_t^T A_t x_t + x_t^T b_t + c_t - (\lambda_t - \lambda_{t-1}) \log \ell(x_{t-1}, y) \\ &: A_t \text{ diagonal}, b_t \in \mathbb{R}^d, c_t \in \mathbb{R} \Big\}, \quad t \in [1 : T] \end{aligned}$$

Log-Gaussian Cox point process

- Parameterization provides good approximation of optimal policy



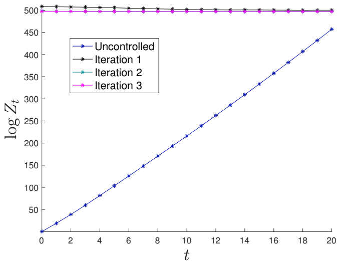


Figure: Effective sample size (*left*) and normalizing constant estimation (*right*) when performing inference on the Scots pine dataset.

Log-Gaussian Cox point process

- Comparison to AIS with MALA moves
- cSMC: $N = 4096$ particles, $I = 3$ iterations, $T = 20$
- AIS uses 5 times more particles for fair comparison
- Variance of marginal likelihood estimates are 280 times smaller

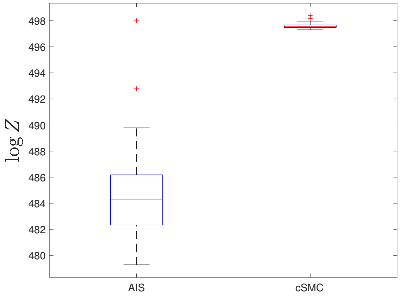
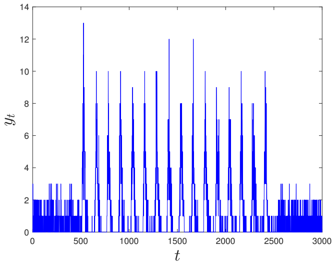


Figure: Marginal likelihood estimates obtained by each algorithm over 100 independent repetitions.

- Measurements collected from a neuroscience experiment (Temereanca et al., 2008)



- State space model:

$$
\begin{aligned}
\mu &= \mathcal{N}(0,1), \\
M_t(x_{t-1}, \mathrm{d}x_t) &= \mathcal{N}\left(x_t; \alpha x_{t-1}, \sigma^2\right) \mathrm{d}x_t, \\
G_t(x_t) &= \mathcal{B}\left(y_t; M, \kappa(x_t)\right),
\end{aligned}
$$

where $M = 30,\ T = 2999$ and $\kappa(u) := (1 + \exp(-u))^{-1}$, for $u \in \mathbb{R}$

A model from neuroscience

- Function classes:
  $\mathsf{F}_t = \left\{\varphi(x_t) = a_t x_t^2 + b_t x_t + c_t : (a_t, b_t, c_t) \in \mathbb{R}^3\right\}, \quad t \in [0 : T],$
- Parameterization provides good approximation of optimal policy

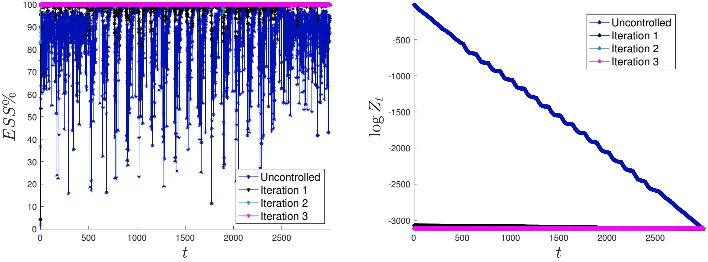


Figure: Effective sample size (*left*) and normalizing constant estimation (*right*) when performing inference on the neuroscience model.

A model from neuroscience

- Estimated policies capturing abrupt changes in the data

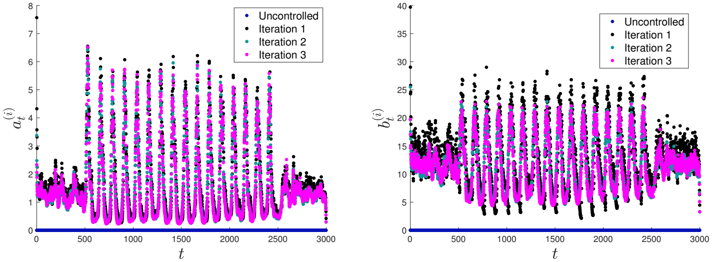


Figure: Coefficients estimated by the controlled SMC sampler at each iteration when performing inference on the neuroscience model.

A model from neuroscience

- (*Left*) Comparison to bootstrap particle filter (BPF)
- (*Right*) Comparison to forward filtering and backward smoother (FFBS) for functional $x_{0:T} \mapsto M\kappa(x_{0:T})$

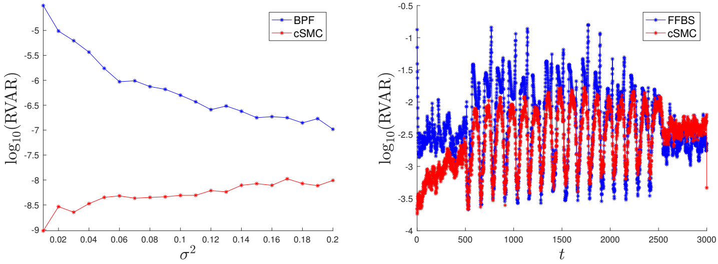


Figure: Relative variance of marginal likelihood estimates (*left*) and estimates of smoothing expectation (*right*).

A model from neuroscience

- Bayesian inference for parameters $\theta = (\alpha, \sigma^2)$ within particle marginal Metropolis-Hastings (PMMH)
- cSMC and BPF to produce unbiased estimates of marginal likelihood

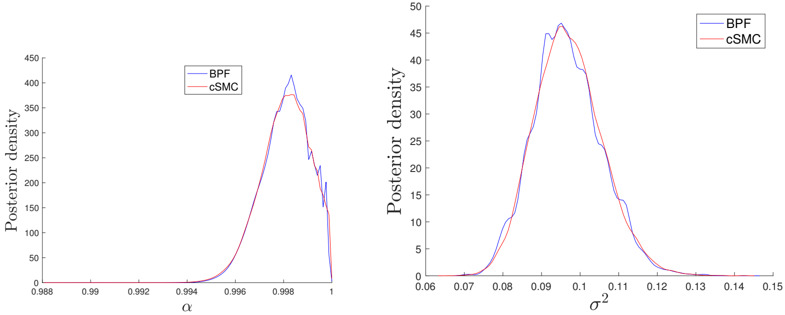


Figure: Posterior density estimates based on 100,000 samples.

A model from neuroscience

- Autocorrelation function (ACF) of each PMMH chain
- ESS improvement roughly 10 times for parameter α and 5 times for parameter σ^2

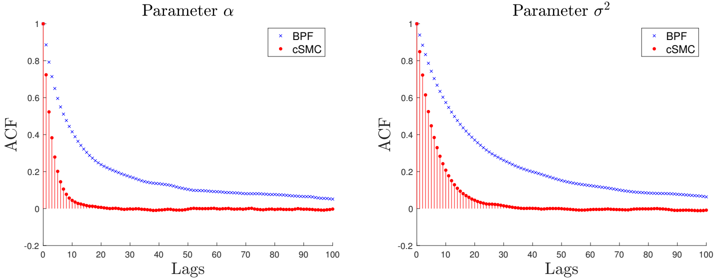


Figure: Autocorrelation functions of two PMMH chains.