Coupling Particle Systems

Pierre E. Jacob (Harvard)
joint work with
Fredrik Lindsten and Thomas Schön (Uppsala)

CRiSM Workshop on Estimating Constants
April 20, 2016

Outline

Hidden Markov models

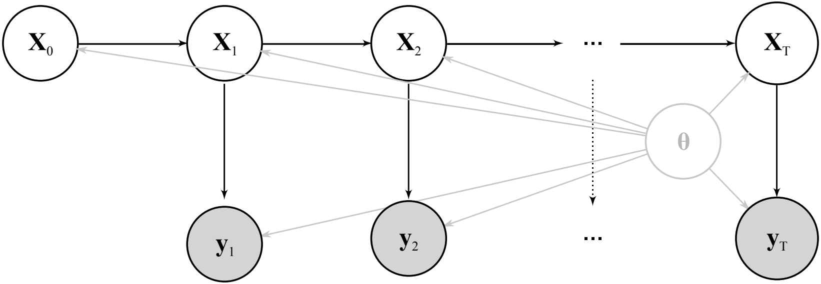


Figure: Graph representation of a general hidden Markov model.

(X_t): initial μ_θ, transition f_θ. (Y_t) given (X_t): measurement g_θ.

Hidden Markov models

- How to estimate/predict the latent process (X_t) given the observations (Y_t) and a fixed parameter θ?

- How to estimate the parameter θ?

Example: Hidden Autoregressive

- Hidden process $X_t = AX_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim \mathcal{N}_d(0, I)$, $X_0 \sim \mathcal{N}_d(0, I)$.
- $A_{ij} = \theta^{|i-j|+1}$ for $i, j \in 1 : d$.
- Observations $Y_t = X_t + \eta_t$, where $\eta_t \sim \mathcal{N}_d(0, I)$.

taken from Guarniero, Johansen & Lee, 2015.

Example: Phytoplankton–Zooplankton

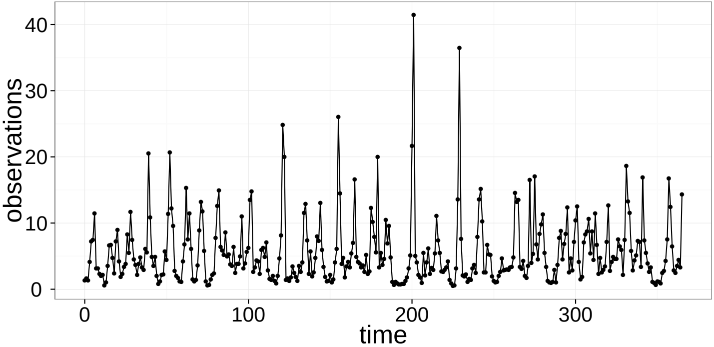


Figure: A time series of 365 observations generated according to a phytoplankton–zooplankton model.

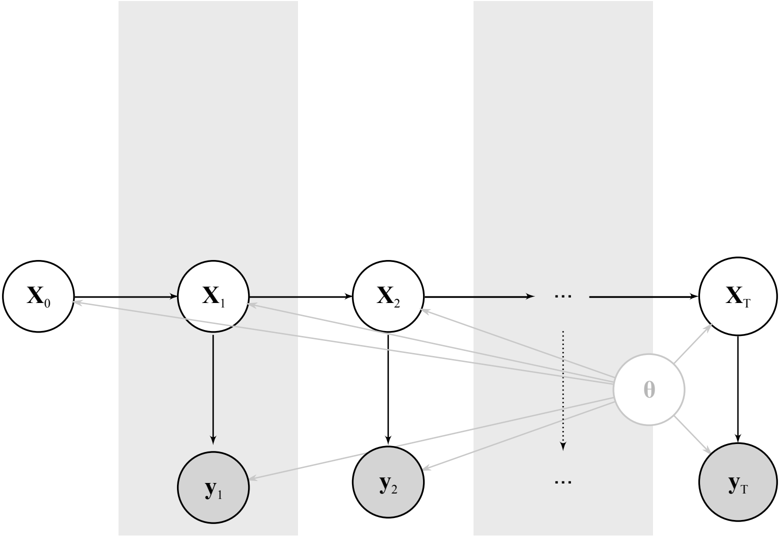
Example: Phytoplankton–Zooplankton

- Hidden process $(X_t) = (\alpha_t, p_t, z_t)$.
- At each (integer) time, $\alpha_t \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$.
- Given α_t,

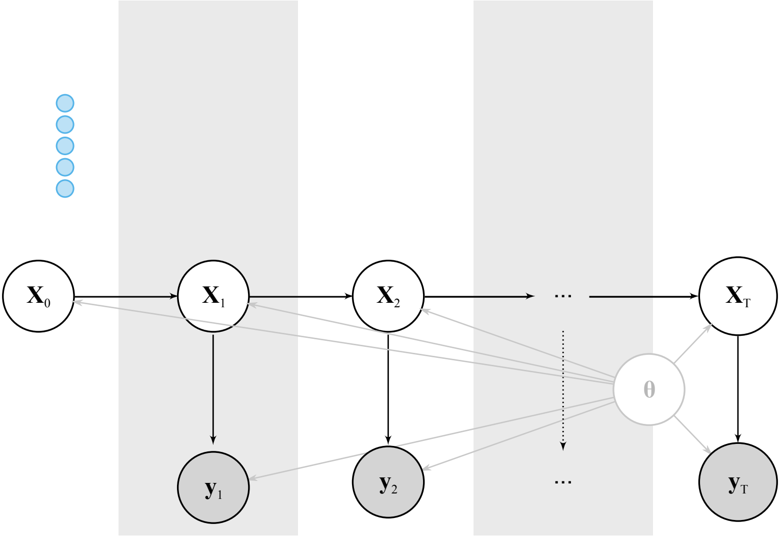
$$\frac{dp_t}{dt} = \alpha_t p_t - c p_t z_t,$$
$$\frac{dz_t}{dt} = e c p_t z_t - m_l z_t - m_q z_t^2.$$

- Observations: $\log Y_t \sim \mathcal{N}(\log p_t, \sigma_y^2)$.
- Initial distribution: $(\log p_0, \log z_0) \sim \mathcal{N}(\mu_0, \sigma_0^2)$.
- Unknown parameters: $\theta = (\mu_0, \sigma_0, \mu_\alpha, \sigma_\alpha, \sigma_y, c, e, m_l, m_q)$.

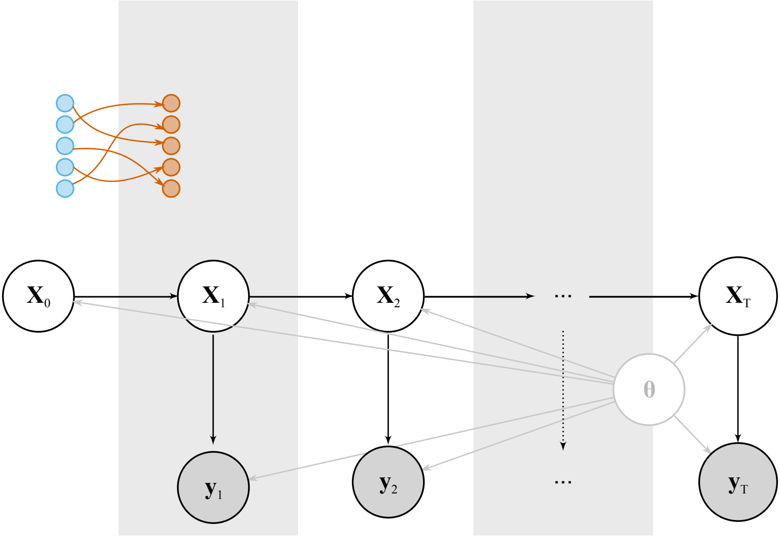
Particle filter



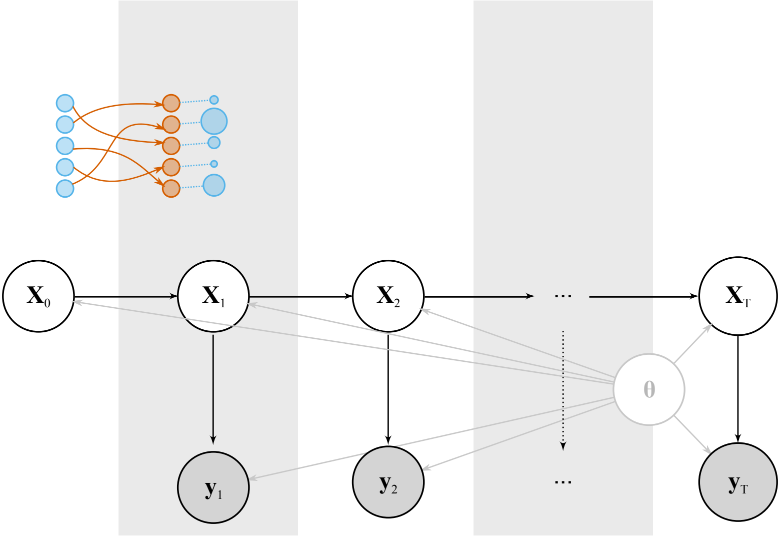
Particle filter



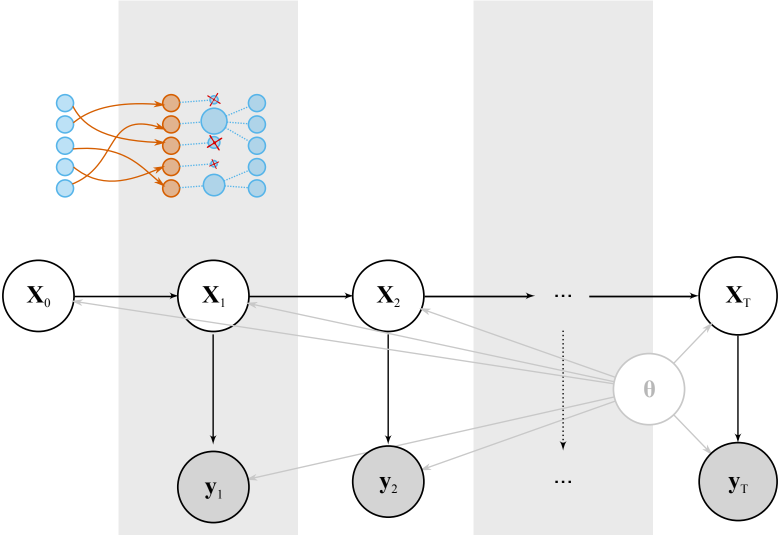
Particle filter



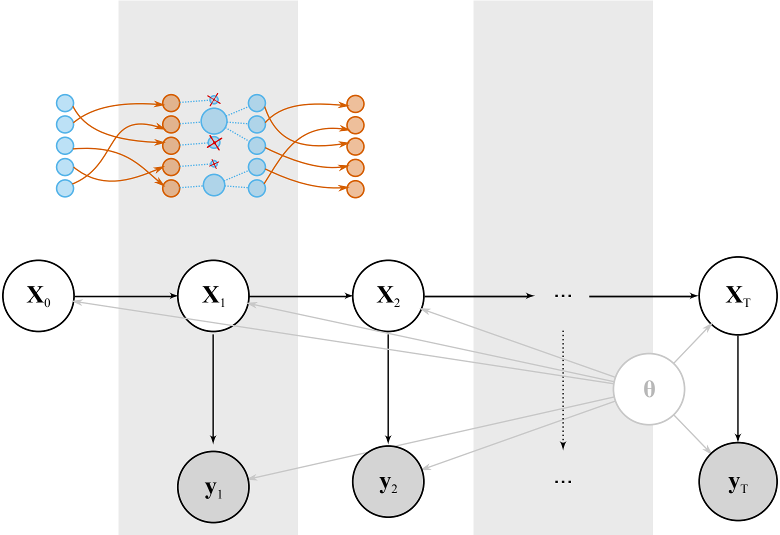
Particle filter



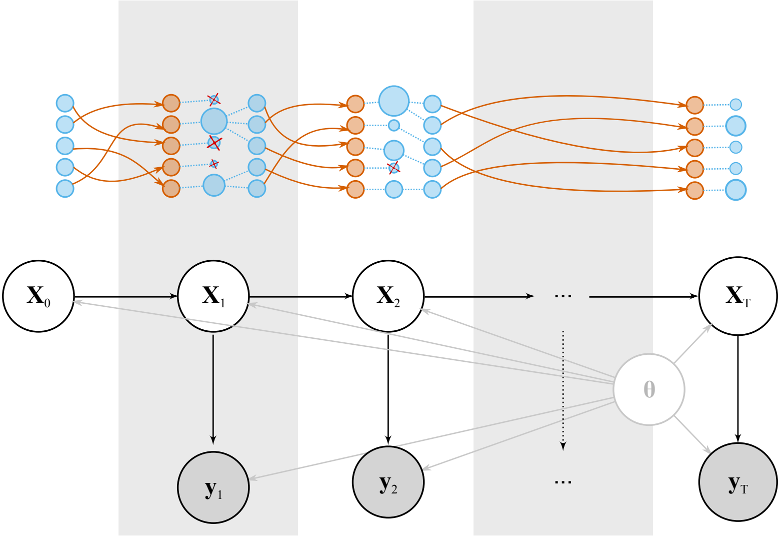
Particle filter



Particle filter



Particle filter



Particle filter

At step $t = 0$,

1. Sample $x_0^k \sim \mu_\theta(dx_0)$, for all $k \in 1:N$.
2. Set $w_0^k = N^{-1}$, for all $k \in 1:N$.

At step $t \geq 1$,

1. Sample ancestors $a_t^{1:N} \sim r(da^{1:N} \mid w_{t-1}^{1:N})$. $\leftarrow$ resampling
2. Sample $x_t^k \sim f_\theta(dx_t \mid x_{t-1}^{a_t^k})$, for all $k \in 1:N$.
3. Compute $w_t^k = g_\theta(y_t \mid x_t^k)$, for all $k \in 1:N$.

Particle filter, rewritten

At step $t = 0$,

1. Sample U_M^k, compute $x_0^k = M(U_M^k, \theta)$, for all $k \in 1 : N$.
2. Set $w_0^k = N^{-1}$, for all $k \in 1 : N$.

At step $t \geq 1$,

1. Sample ancestors $a_t^{1:N} \sim r(da^{1:N} \mid w_{t-1}^{1:N})$. $\leftarrow$ resampling
2. Sample $U_{F,t}^k$, compute $x_t^k = F(x_{t-1}^{a_t^k}, U_{F,t}^k, \theta)$, for all $k \in 1 : N$.
3. Compute $w_t^k = g_\theta(y_t \mid x_t^k)$, for all $k \in 1 : N$.

Output

- Approximation of the filtering distributions

$$\forall t \in \{1, \ldots, T\} \quad p(dx_t | y_{1:t}, \theta)$$

by

$$\forall t \in \{1, \ldots, T\} \quad p^N(dx_t | y_{1:t}, \theta) = \sum_{k=1}^{N} w_t^k \delta_{x_t^k}(dx_t).$$

- Approximation of the likelihood function $\mathcal{L}(\theta) = p(y_{1:T}|\theta)$ by

$$p^N(y_{1:T}|\theta) = \prod_{t=1}^{T} \frac{1}{N} \sum_{k=1}^{N} w_t^k.$$

Idea

- Particle filters are increasingly used as parts of encompassing algorithms.
  e.g. Particle MCMC, Iterated Filtering

- Some of these algorithms *compare* the outputs of multiple particle filters.

- Better algorithms can be obtained by *correlating* particle filters.

i.e. correlation helps comparison.

Example: approximation of the likelihood

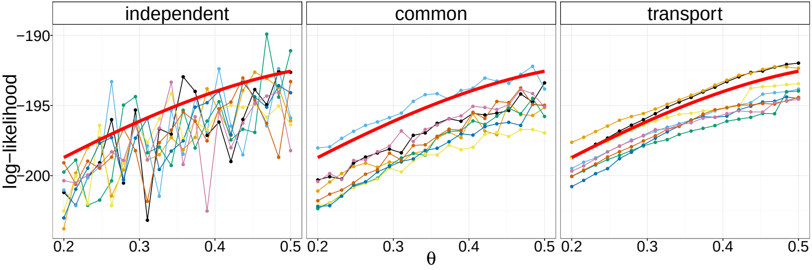


Figure: Estimates of the log-likelihood obtained by particle filters, in a hidden auto-regressive model, $T = 100$ observations, $N = 64$ particles.

See Pitt & Malik, 2011, Lee 2008.

Example: finite difference

A simple estimator of $\nabla \ell(\theta) = \nabla \log \mathcal{L}(\theta)$ is:

$$\widehat{\nabla} \ell(\theta) = \frac{\log \hat{p}^N(y_{1:T} \mid \theta + h) - \log \hat{p}^N(y_{1:T} \mid \theta - h)}{2h}.$$

The two log-likelihood estimators can be obtained using independent particle filters given $\theta + h$ and $\theta - h$...

...but if we could positively correlate the two log-likelihood estimators, the variance of $\widehat{\nabla} \ell(\theta)$ would be smaller.

Example: finite difference

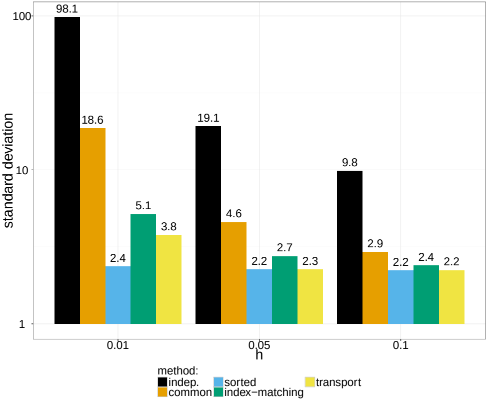


Figure: Standard deviation of $\widehat{\nabla \ell}(\theta)$, for some θ, in a hidden auto-regressive model, $T = 100$ observations, $N = 128$ particles.

Example: finite difference

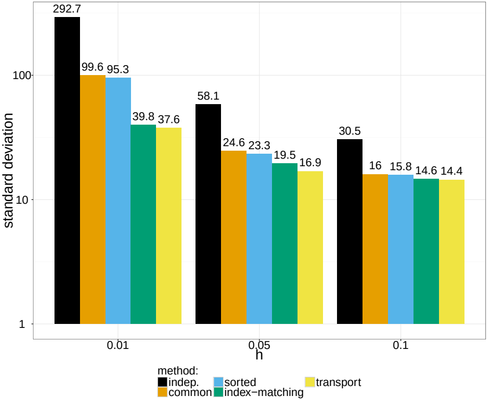


Figure: Same but in dimension 5.

Example: finite difference

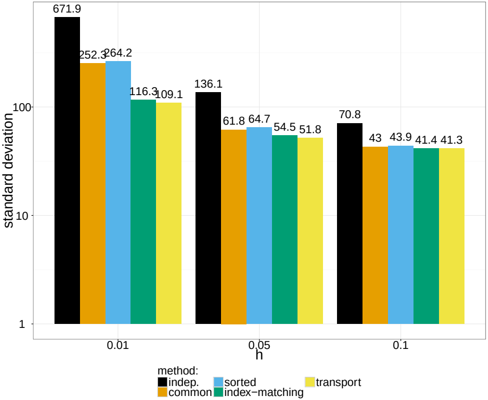


Figure: Same but in dimension 10.

Example: Metropolis-Hastings

Assume we can compute the target density $\pi(\theta|y_{1:T})$ pointwise.

1: Set some $\theta^{(1)}$.
2: **for** $i = 2$ to M **do**
3: Propose $\theta^\star \sim q(\cdot|\theta^{(i-1)})$.
4: Compute the ratio:

$$\alpha = \min\left(1, \frac{\pi(\theta^\star)p(y_{1:T}|\theta^\star)}{\pi(\theta^{(i-1)})p(y_{1:T}|\theta^{(i-1)})}\frac{q(\theta^{(i-1)}|\theta^\star)}{q(\theta^\star|\theta^{(i-1)})}\right).$$

5: Set $\theta^{(i)} = \theta^\star$ with probability α, otherwise set $\theta^{(i)} = \theta^{(i-1)}$.
6: **end for**

Example: particle Metropolis-Hastings

Assume we can run a particle filter to get $p^N(y_{1:T}|\theta)$.

1: Set some $\theta^{(1)}$ and sample $p^N(y_{1:T}|\theta^{(1)})$.
2: **for** $i = 2$ to M **do**
3: Propose $\theta^\star \sim q(\cdot|\theta^{(i-1)})$ and sample $p^N(y_{1:T}|\theta^\star)$.
4: Compute the ratio:

$$\alpha = \min\left(1, \frac{\pi(\theta^\star)p^N(y_{1:T}|\theta^\star)}{\pi(\theta^{(i-1)})p^N(y_{1:T}|\theta^{(i-1)})}\frac{q(\theta^{(i-1)}|\theta^\star)}{q(\theta^\star|\theta^{(i-1)})}\right).$$

5: Set $\theta^{(i)} = \theta^\star$ with probability α, otherwise set $\theta^{(i)} = \theta^{(i-1)}$.
6: **end for**

Andrieu, Doucet & Holenstein, 2010.

Example: particle Metropolis-Hastings

- The acceptance ratio involves a ratio of particle filter estimators:

$$\alpha = \min\left(1, \frac{\pi(\theta^\star) p^N(y_{1:T}|\theta^\star)}{\pi(\theta^{(i-1)}) p^N(y_{1:T}|\theta^{(i-1)})} \frac{q(\theta^{(i-1)}|\theta^\star)}{q(\theta^\star|\theta^{(i-1)})}\right).$$

- If we positively correlate $p^N(y_{1:T}|\theta^\star)$ and $p^N(y_{1:T}|\theta^{(i-1)})$, the ratio of estimators becomes more precise.

Deligiannidis, Doucet, Pitt & Kohn, 2015: epic improvements for the case large T / small N.

Outline

Coupled particle filter

By coupled particle filters we mean ...

- two particle systems, $(w_t^k, x_t^k)_{k=1}^N$ and $(\tilde{w}_t^k, \tilde{x}_t^k)_{k=1}^N$,
- conditioned on θ and $\tilde{\theta}$ respectively,
- using common random numbers U_M and U_F for the initial and propagation steps.

One still has the freedom to choose a "coupled resampling" scheme.

Coupled particle filter

At step $t = 0$,

1. Sample U_M^k, and compute, for all $k \in 1:N$,

$$x_0^k = M(U_M^k, \theta) \text{ and } \tilde{x}_0^k = M(U_M^k, \tilde{\theta}).$$

2. Set $w_0^k = N^{-1}$ and $\tilde{w}_0^k = N^{-1}$, for all $k \in 1:N$.

At step $t \geq 1$,

1. Sample ancestors:

$$(a_t, \tilde{a}_t) \sim \bar{r}(\cdot | w_{t-1}^{1:N}, \tilde{w}_{t-1}^{1:N}). \qquad \leftarrow \text{coupled resampling}$$

2. Sample $U_{F,t}^k$, and compute, for all $k \in 1:N$,

$$x_t^k = F(x_{t-1}^{a_t^k}, U_{F,t}^k, \theta) \text{ and } \tilde{x}_t^k = F(\tilde{x}_{t-1}^{\tilde{a}_t^k}, U_{F,t}^k, \tilde{\theta}).$$

3. Compute $w_t^k = g(y_t \mid x_t^k, \theta)$ and $\tilde{w}_t^k = g(y_t \mid \tilde{x}_t^k, \tilde{\theta})$.

Coupled resampling

Given two particle systems, $(w^k, x^k)_{k=1}^N$ and $(\tilde{w}^k, \tilde{x}^k)_{k=1}^N \ldots$

- We want (?) to sample $a^{1:N}$ and $\tilde{a}^{1:N}$ in $\{1, \ldots, N\}^N$ such that

$$\forall k \quad \forall j \quad \mathbb{P}(a^k = j) = w^j \text{ and } \mathbb{P}(\tilde{a}^k = j) = \tilde{w}^j.$$

- Equivalently, we want to sample $(a^k, \tilde{a}^k)_{k=1}^N$ from a probability matrix P such that

$$P\mathbb{1} = w \quad \text{and} \quad P^T\mathbb{1} = \tilde{w}.$$

Independent resampling corresponds to $P = w\,\tilde{w}^T$. What else?

Transport resampling

- Suppose that we want to sample a couple $(a, \tilde{a})$, from some probability matrix P, such that the resampled particles, x^a and $\tilde{x}^{\tilde{a}}$, are as similar as possible.

- Similarity can be encoded by a distance d on the space of x.

- The expected distance between x^a and $\tilde{x}^{\tilde{a}}$, conditional upon the particles, is given by

$$\mathbb{E}\left[d(x^a, \tilde{x}^{\tilde{a}})\right] = \sum_{i=1}^{N}\sum_{j=1}^{N} P_{ij}\, d(x^i, \tilde{x}^j).$$

Transport resampling

- Introduce $\mathcal{J}(w, \tilde{w})$, the set of matrices satisfying

$$P\mathbb{1} = w \quad \text{and} \quad P^T\mathbb{1} = \tilde{w}.$$

- Compute $D = (d(x^i, \tilde{x}^j))_{i,j=1}^N$, for a cost of $\mathcal{O}(N^2)$.

- Optimal transport problem: solving

$$P^\star = \inf_{P\in\mathcal{J}(w,\tilde{w})} \sum_{i=1}^N \sum_{j=1}^N P_{ij} D_{ij}.$$

Transport resampling

- Sampling from the optimal $P^\star$ minimizes the expected distance between the two sets of particles, under the marginal constraint.
- (Which is not exactly the same as maximizing the correlation between e.g. likelihood estimators).
- Computing $P^\star$ requires $\mathcal{O}(N^3)$ operations, but efficient approximations have been proposed (Cuturi 2013 and following work) in $\mathcal{O}(N^2)$.
- The cost is linear in the dimension of x, and independent of the model.

Index-matching resampling

- At the initial step,

$$x_0^k = M(U_M^k, \theta) \text{ and } \tilde{x}_0^k = M(U_M^k, \tilde{\theta}).$$

so that particles with the same index are similar (if M is continuous in θ).

- At subsequent steps, the same random numbers U_F^k are used to propagate x^{a^k} and $\tilde{x}^{\tilde{a}^k}$.

- Standard resampling breaks the correspondence between similarity and indices: x^{a^k} and $\tilde{x}^{\tilde{a}^k}$ might not be similar.

Index-matching resampling

- To preserve the correspondence between indices, we want to maximize the probability of drawing couples of ancestors such that $a = \tilde{a}$.
- ... i.e. we want P in $\mathcal{J}(w, \tilde{w})$ with maximum trace.
- We can define:

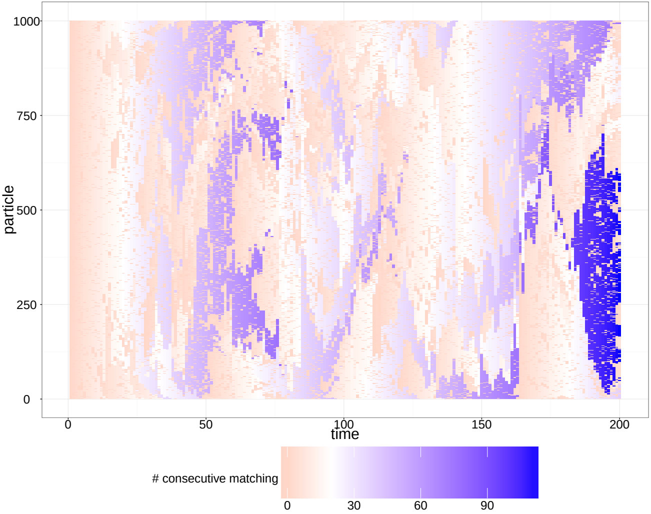
$$P = \mathrm{diag}(\min\{w, \tilde{w}\}) + (1-\alpha) r \, \tilde{r}^T,$$

with

$$\alpha = \sum_{k=1}^{N} \min\{w^k, \tilde{w}^k\},$$
$$r = (w - \min\{w, \tilde{w}\})/(1-\alpha),$$
$$\tilde{r} = (\tilde{w} - \min\{w, \tilde{w}\})/(1-\alpha).$$

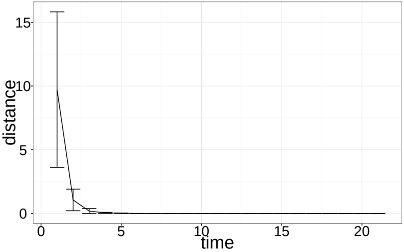
- P has maximum trace in $\mathcal{J}(w, \tilde{w})$: we cannot augment its diagonal without violating the marginal constraints.

Index-matching resampling



Index-matching resampling

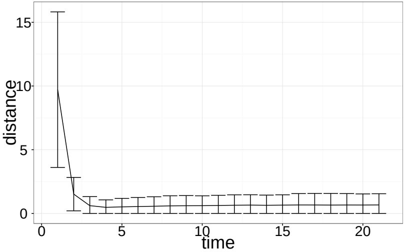
Distance between $1,000$ pairs of particles, sampled independently and then propagated with common random numbers.



Hidden AR, $\theta = 0.30$, $\tilde{\theta} = 0.31$.

Index-matching resampling

Distance between $1,000$ pairs of particles, sampled independently and then propagated with common random numbers.



Hidden AR, $\theta = 0.30$, $\tilde{\theta} = 0.40$.

Sorting and resampling

Univariate setting: x is of dimension 1.

- Sort the two systems $(x^k)_{k=1}^N$ and $(\tilde{x}^k)_{k=1}^N$.
- Perform e.g. systematic resampling on each sorted system, using the same random numbers.
- Thus if a^k selects x^j in the first system, $\tilde{a}^k$ is likely to select a $\tilde{x}^i$ close to x^j.

This can be extended to multivariate settings by sorting the particles according to the Hilbert space-filling curve.
See Deligiannidis, Doucet, Pitt & Kohn, 2015, and Pitt & Malik, 2011, Gerber & Chopin, 2015.

Outline

Who cares about unbiased estimators?

- Coupled resampling leads to a practical unbiased estimator H_u of the smoothing quantity

$$\int h(x_{0:T})\, p(dx_{0:T}|y_{1:T}, \theta),$$

for a test function h (for fixed θ).

- Computing $H_u^{(1)}, \ldots, H_u^{(R)}$ in parallel, we obtain

$$\bar{H}_u = \frac{1}{R}\sum_{r=1}^{R} H_u^{(r)},$$

along with a CLT-based error estimate.

- By contrast, for existing smoothing techniques, parallelism is not trivial, nor is the construction of error estimates.

Proposed estimator

- Use Rhee & Glynn (2014) trick to turn a Conditional Particle Filter kernel into an unbiased estimator of smoothing functionals.
- Coupled resampling schemes are instrumental in this construction.
- Instead of two particle systems given θ and $\tilde{\theta}$, we consider two particle systems with same θ but different "reference trajectories".

Trajectories from particle filters

Upon running a particle filter, we get trajectories $x_{0:T}^{1:N}$ with weights $w_T^{1:N}$.

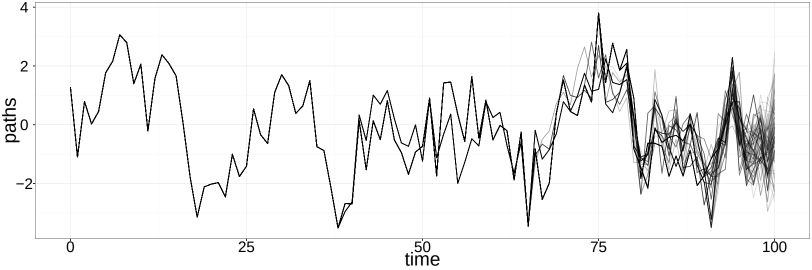


Figure: Hidden auto-regressive model, $T = 100$ observations, $N = 128$.

Conditional particle filter

Input: a trajectory $\check{x}_{0:T}$.
At step $t = 0$,

1. Sample $x_0^k \sim \mu_\theta(dx_0)$, for all $k \in 1:N-1$, set $x_0^N = \check{x}_0$.
2. Set $w_0^k = N^{-1}$, for all $k \in 1:N$.

At step $t \geq 1$,

1. Sample ancestors $a_t^{1:N-1} \sim r(da^{1:N-1} \mid w_{t-1}^{1:N})$, set $a_t^N = N$.
2. Sample $x_t^k \sim f_\theta(dx_t \mid x_{t-1}^{a_t^k})$, for all $k \in 1:N-1$, set $x_t^N = \check{x}_t$.
3. Compute $w_t^k = g_\theta(y_t \mid x_t^k)$, for all $k \in 1:N$.

Output: sample a trajectory, $x_{0:T}^k$ with probability w_T^k.

Conditional particle filter

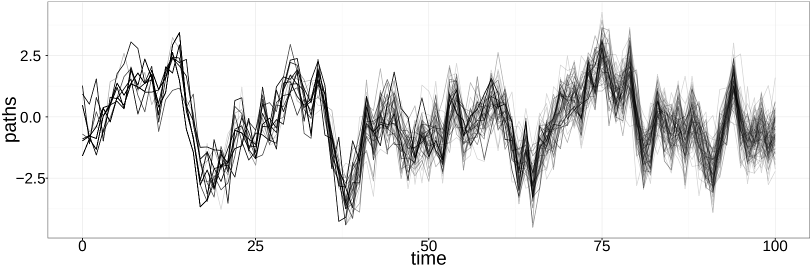


Figure: $M = 100$ paths, for the hidden auto-regressive model, $T = 100$ observations, $N = 128$.

Conditional particle Filter

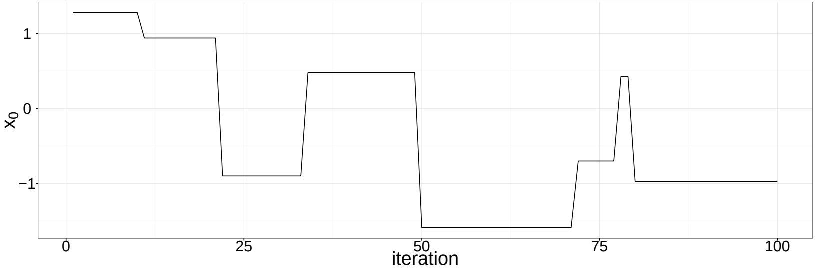


Figure: $M = 100$ samples for x_0, for the hidden auto-regressive model, $T = 100$ observations, $N = 128$.

Conditional particle Filter

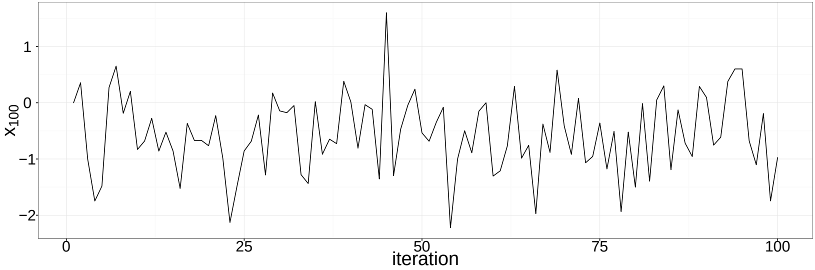


Figure: $M = 100$ samples for x_{100}, for the hidden auto-regressive model, $T = 100$ observations, $N = 128$.

Unbiased estimators

Von Neumann & Ulam (~ 1950), Kuti (~ 1980), Rychlik (~ 1990), McLeish (~ 2010), **Rhee & Glynn** $(2012, 2013, 2014)$.

Introduce

- a sequence of random variables $(H^{(n)})$ with

$$\mathbb{E}[H^{(n)}] \xrightarrow[n\to\infty]{} \int h(x_{0:T})\, p(dx_{0:T}|y_{1:T},\theta),$$

  e.g. $H^{(n)} = h(X^{(n)})$, with $(X^{(n)})$ generated by CPF,
- a sequence $(\Delta^{(n)})$ such that

$$\mathbb{E}[\Delta^{(n)}] = \mathbb{E}[H^{(n)} - H^{(n-1)}],$$

$$\mathbb{E}\left[\sum_{n=0}^{\infty} |\Delta^{(n)}|\right] < \infty,$$

  with $H^{(-1)} = 0$ by convention.

Unbiased estimators

- Then

$$\mathbb{E}\sum_{n=0}^{\infty}\Delta^{(n)} = \sum_{n=0}^{\infty}\mathbb{E}[\Delta^{(n)}] = \sum_{n=0}^{\infty}\mathbb{E}[H^{(n)} - H^{(n-1)}]$$
$$= \lim_{n\to\infty}\mathbb{E}[H^{(n)}] = \int h(x_{0:T})\, p(dx_{0:T}|y_{1:T},\theta).$$

Thus, consider

$$H_u = \sum_{n=0}^{K}\frac{\Delta^{(n)}}{\mathbb{P}\left(K \geq n\right)},$$

where K is an integer-valued random variable. Then

$$\mathbb{E}[H_u] = \mathbb{E}[\sum_{n=0}^{\infty}\frac{\Delta^{(n)}\mathbb{1}(K\geq n)}{\mathbb{P}\left(K\geq n\right)}] = \int h(x_{0:T})\, p(dx_{0:T}|y_{1:T},\theta).$$

Unbiased estimators

Idea from Rhee & Glynn, 2014. Write

$$X^{(n)} = \varphi_n(X^{(n-1)}) = \varphi_n \circ \varphi_{n-1} \circ \ldots \circ \varphi_1(X^{(0)}).$$

Introduce

$$\tilde{X}^{(0)} \triangleq X^{(0)}, \quad \tilde{X}^{(1)} = \varphi_2(\tilde{X}^{(0)}), \quad \ldots, \quad \tilde{X}^{(n)} = \varphi_{n+1} \circ \ldots \circ \varphi_2(X^{(0)}).$$

Then $\Delta^{(n)} = h(X^{(n)}) - h(\tilde{X}^{(n-1)})$ is such that

$$\mathbb{E}[\Delta^{(n)}] = \mathbb{E}[H^{(n)} - H^{(n-1)}]$$

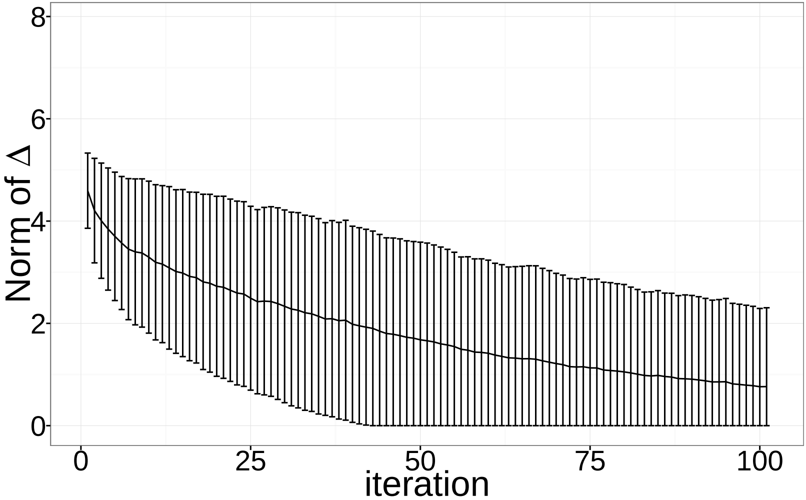
and we might have

$$\mathbb{E}\left[\sum_{n=0}^{\infty} |\Delta^{(n)}|\right] < \infty.$$

Unbiased estimators based on CPF chains

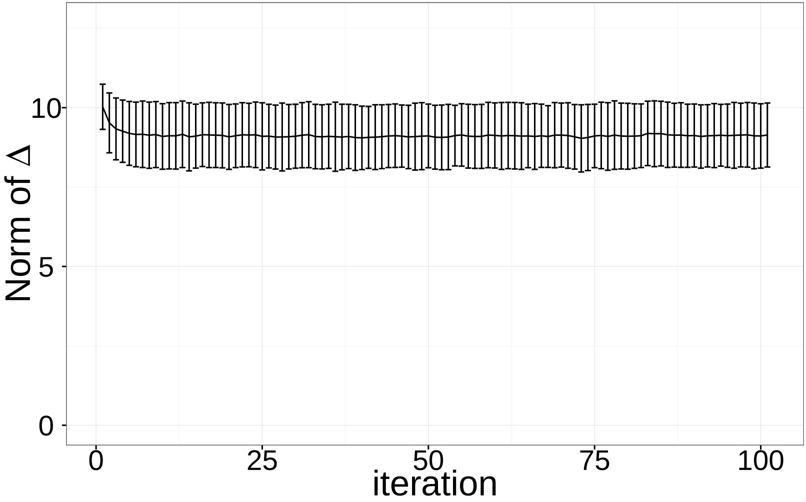
- Start from $X^{(0)}$ and $\tilde{X}^{(0)}$ generated by two particle filters.
- Apply one step of CPF kernel to $X^{(0)}$, to get $X^{(1)}$.
- For $n \geq 2$, apply the CPF kernel to both $X^{(n-1)}$ and $\tilde{X}^{(n-2)}$, with the same random numbers, to get $X^{(n)}$ and $\tilde{X}^{(n-1)}$.
- We can see each step as a joint CPF acting on pairs of trajectories, and use coupled resampling ideas.
- Can we expect $\Delta^{(n)} = h(X^{(n)}) - h(\tilde{X}^{(n-1)})$ to decrease to zero in average?

Norm of $\Delta^{(n)}$ with independent resampling



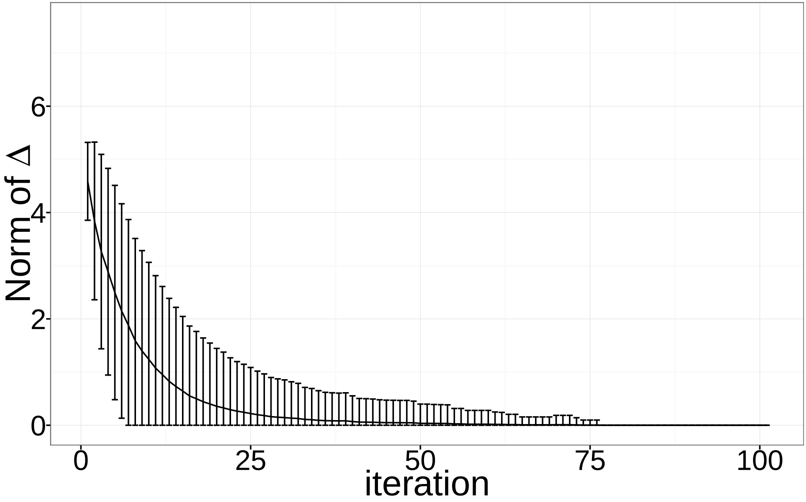
Hidden auto-regressive model, $T = 20$ observations, $N = 32$.

Norm of $\Delta^{(n)}$ with independent resampling



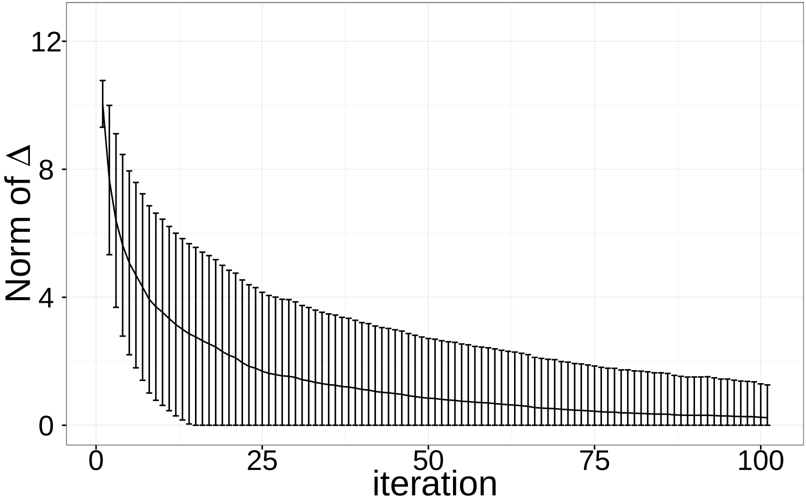
$T = 100$ observations, $N = 128$.

Norm of $\Delta^{(n)}$ with index-matching resampling



$T = 20$ observations, $N = 32$.

Norm of $\Delta^{(n)}$ with index-matching resampling



$T = 100$ observations, $N = 128$.

Coupled conditional particle Filter

- We consider coupled conditional particle filters, acting on pairs of trajectories:

$$(X^{(n)}, \tilde{X}^{(n-1)}) = \bar{\varphi}_n(X^{(n-1)}, \tilde{X}^{(n-2)})$$

- A coupled CPF kernel uses common random numbers for both systems, and a coupled resampling scheme.
- We focus on index-matching resampling.
- We see that after a number of coupled CPF steps, $X^{(n)} = \tilde{X}^{(n)}$ exactly, and thus $\Delta^{(n)} = 0$.
- We can thus stop early in the computation of

$$H_u = \sum_{n=0}^{K} \frac{\Delta^{(n)}}{\mathbb{P}\left(K \geq n\right)}.$$

Proposed estimator

Case $h = Id$: we estimate the smoothing means.

- Sample an integer-valued random variable K.
- Sample φ_1, draw $X^{(0)}$ and set $X^{(1)} = \varphi_1(X^{(0)})$.
- Compute $\Delta^{(0)} = X^{(0)}$, set $H_u \leftarrow \Delta^{(0)}$.
- Sample $\tilde{X}^{(0)} \stackrel{\Delta}{=} X^{(0)}$, compute $\Delta^{(1)} = X^{(1)} - \tilde{X}^{(0)}$.
- Set $H_u \leftarrow H_u + \Delta^{(1)}/\mathbb{P}(K \geq 1)$.
- For $n = 2, \ldots, K$,
  - Sample $\bar{\varphi}_n$, set $(X^{(n)}, \tilde{X}^{(n-1)}) = \bar{\varphi}_n(X^{(n-1)}, \tilde{X}^{(n-2)})$.
  - Compute $\Delta^{(n)} = X^{(n)} - \tilde{X}^{(n-1)}$.
  - Stop if $\Delta^{(n)} = 0$.
  - Set $H_u \leftarrow H_u + \Delta^{(n)}/\mathbb{P}(K \geq n)$.
- Return H_u.

Example: Phytoplankton–Zooplankton

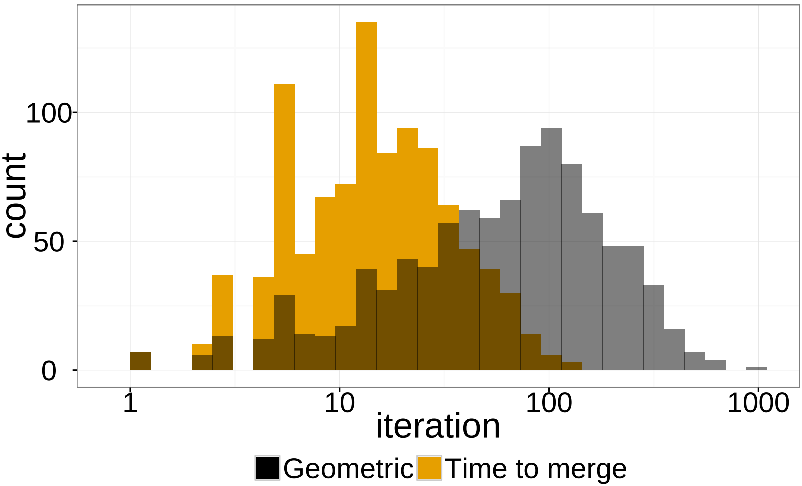


Figure: Phytoplankton–Zooplankton model, $T = 365$, $N = 1,024$, $R = 1,000$ estimators, with a Geometric truncation with mean 100.

Example: Phytoplankton–Zooplankton

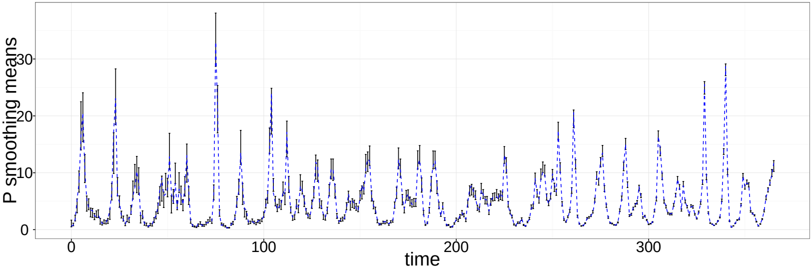


Figure: Smoothing means of P, $T = 365$, $N = 1,024$, $R = 1,000$ estimators, with a Geometric truncation with mean 100.

The bars represent $\pm 2\sigma$ around the estimated means. The blue line is obtained from a long CPF run.

Example: Phytoplankton–Zooplankton

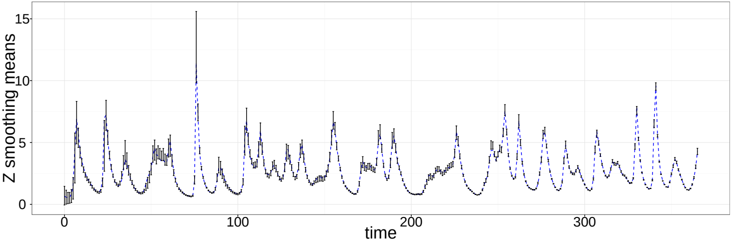


Figure: Smoothing means of Z, $T = 365$, $N = 1,024$, $R = 1,000$ estimators, with a Geometric truncation with mean 100.

The bars represent $\pm 2\sigma$ around the estimated means. The blue line is obtained from a long CPF run.

Example: Phytoplankton–Zooplankton

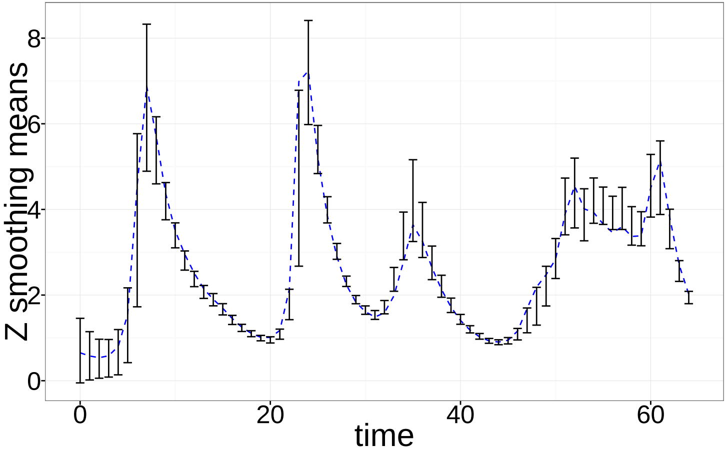


Figure: Smoothing means of Z, first 65/365 time steps, $N = 1,024$, $R = 1,000$ estimators, with a Geometric truncation with mean 100.

Example: Phytoplankton–Zooplankton

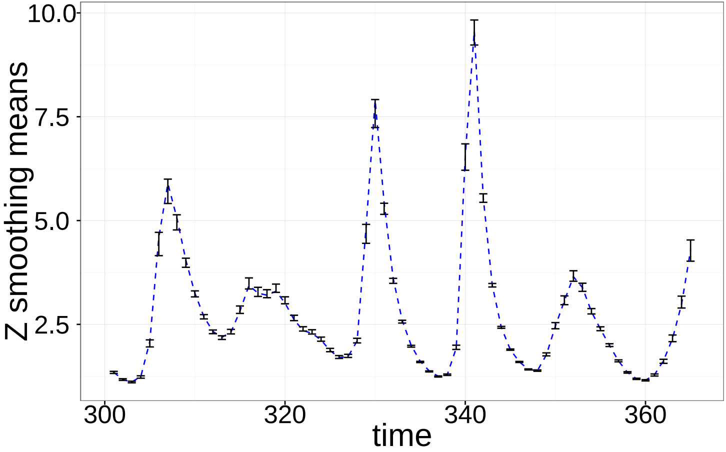


Figure: Smoothing means of Z, last 65/365 time steps, $N = 1,024$, $R = 1,000$ estimators, with a Geometric truncation with mean 100.

Example: Phytoplankton–Zooplankton

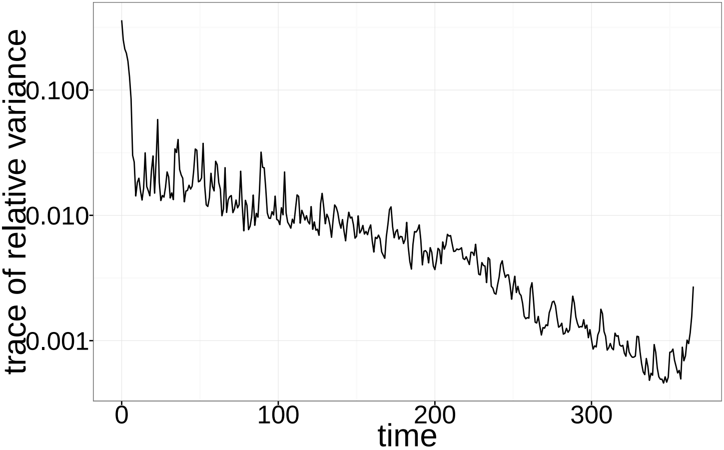
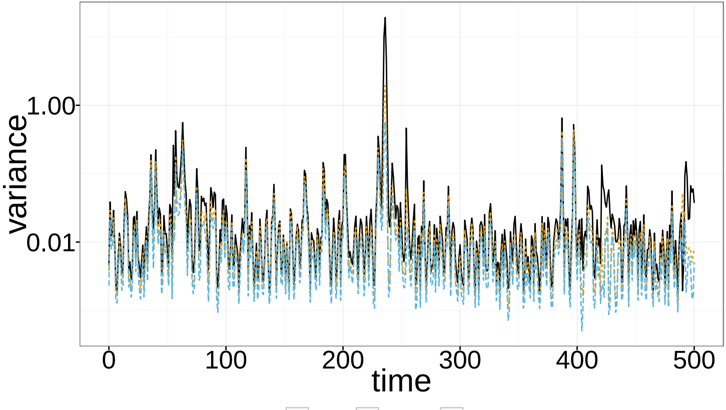


Figure: Trace of relative variance of the smoothing mean estimator, $T = 365$, $N = 1,024$, $R = 1,000$ estimators, with a Geometric truncation with mean 100.

Discussion

- Coupled resampling schemes can be used to improve a variety of particle-based algorithms.
- New estimator of smoothing functionals, easy to parallelize and with error estimates.
- Benefits greatly from ancestor sampling:



The End

Thank you for listening!

Soon on arXiv...
PJ, Fredrik Lindsten, Thomas Schön, *Coupling Particle Filters.*

- Pitt & Malik, 2011, *Particle filters for continuous likelihood evaluation and maximisation*, J. of Econometrics.
- Rhee & Glynn, 2014, *Exact estimation for markov chain equilibrium expectations*, arXiv.
- Deligiannidis, Doucet, Pitt & Kohn, 2015, *The correlated pseudo-marginal method*, arXiv.