Block-Tempered Particle Filters

Adam M. Johansen

University of Warwick
a.m.johansen@warwick.ac.uk
www2.warwick.ac.uk/fac/sci/statistics/staff/academic/johansen/talks/

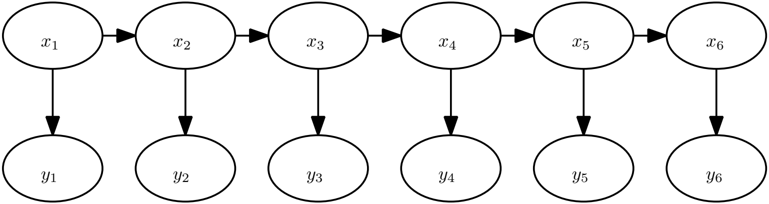
SysID 17: October 20th, 2015

Outline

- Background: SMC and PMCMC
- Tempering Blocks in SMC
- Example: proof-of-concept results
- Conclusions

Filtering

Online inference for State Space Models:



- Given *transition* $f_\theta(x_n|x_{n-1})$,
- and *likelihood* $g_\theta(y_n|x_n)$,
- use $p_\theta(x_n|y_{1:n})$ to characterize latent state, but,

$$p_\theta(x_n|y_{1:n}) = \frac{\int p_\theta(x_{n-1}|y_{1:n-1}) f_\theta(x_n|x_{n-1}) dx_{n-1} g_\theta(y_n|x_n)}{\int \int p_\theta(x_{n-1}|y_{1:n-1}) f_\theta(x'_n|x_{n-1}) dx_{n-1} g_\theta(y_n|x'_n) dx'_n}$$

isn't often tractable.

Particle Filtering

A (sequential) Monte Carlo (SMC) scheme to approximate the filtering distributions.

A Simple Particle Filter

At $n = 1$:

- Sample $X_1^1, \dots, X_1^N \sim \mu_\theta$.

For $n > 1$:

- Sample

$$X_n^1, \dots, X_n^N \sim \frac{\sum_{j=1}^N g_\theta(y_{n-1}|X_{n-1}^j) f_\theta(\cdot|X_{n-1}^j)}{\sum_{k=1}^n g_\theta(y_{n-1}|X_{n-1}^k)}$$

- Approximate $p_\theta(dx_n|y_{1:n}), p_\theta(y_{1:n})$ with

$$\widehat{p_\theta}(\cdot|y_{1:n}) = \frac{\sum_{j=1}^N g_\theta(y_n|X_n^j)\delta_{X_n^j}}{\sum_{k=1}^n g_\theta(y_n|X_n^k)}, \frac{\widehat{p_\theta}(y_{1:n})}{\widehat{p_\theta}(y_{1:n-1})} = \frac{1}{n}\sum_{j=1}^N g_\theta(y_n|X_n^k)$$

Online Particle Filters for Offline Systems Identification

Particle Markov chain Monte Carlo (PMCMC) [ADH10]

- Embed SMC within MCMC,
- justified via explicit auxiliary variable construction,
- and in some simple cases by a pseudomarginal [AR09] argument.
- Very widely applicable,
- but prone to poor mixing when SMC performs poorly for some θ [OWG15, Section 4.2.1].
- Is valid for *very* general SMC algorithms.

Block-Sampling and Tempering in Particle Filters

Tempered Transitions [GC01]

- Introduce each likelihood term gradually, targetting:

$$\pi_{n,m}(x_{1:n}) \propto p(x_{1:n}|y_{1:n-1})p(y_n|x_n)^{\beta_m}$$

  between $p(x_{1:n-1}|y_{1:n-1})$ and $p(x_{1:n}|y_{1:n})$.
- Can improve performance — but to a limited extent.

Block Sampling [DBS06]

- Essentially uses $y_{n:n+L}$ in proposing x_n.
- Can *dramatically* improve performance,
- but requires good analytic approximation of $p(x_{n:n+L}|x_{n-1}, y_{n:n+L})$.

Block-Tempering

- We could combine blocking and tempering strategies.
- Run a simple SMC sampler [DDJ06] targetting:

$$\pi_{t,r}^{\theta}(x_{1:t\wedge T}) = \mu_\theta(x_1)^{\beta_{(t,r)}^1} g_\theta(y_1|x_1)^{\gamma_{(t,r)}^1} \cdot \prod_{s=2}^{T\wedge t} f_\theta(x_s|x_{s-1})^{\beta_{(t,r)}^s} g_\theta(y_s|x_s)^{\gamma_{(t,r)}^s}, \quad (1)$$

  where $\{\beta_{(t,r)}^s\}$ and $\{\gamma_{(t,r)}^s\}$ are $[0,1]$-valued
- for $s \in [\![1,T]\!]$, $r \in [\![1,R]\!]$ and $t \in [\![1,T']\!]$, with $T' = T + L$
- for some $R, L \in \mathbb{N}$.
- Can be *validly* embedded within PMCMC:
  - Terminal likelihood estimate is unbiased.
  - Explicit auxiliary variable construction is possible.

Two Simple Block-Tempering Strategies

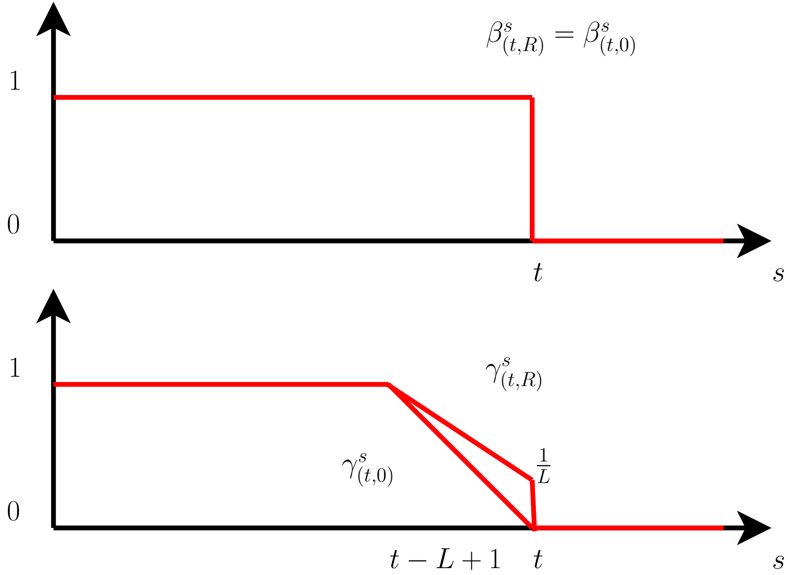
Tempering both likelihood and transition probabilities

$$\beta^s_{(t,r)} = \gamma^s_{(t,r)} = \left(1 \wedge \frac{R(t-s)+r}{RL}\right) \vee 0 \tag{2}$$

Tempering only the observation density

$$\beta^s_{(t,r)} = \mathbb{I}\{s \leq t\} \qquad \gamma^s_{(t,r)} = \left(1 \wedge \frac{R(t-s)+r}{RL}\right) \vee 0, \tag{3}$$

Tempering Only the Observation Density

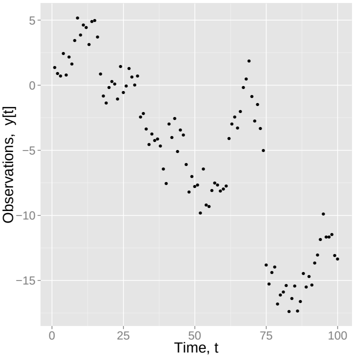


Illustrative Example

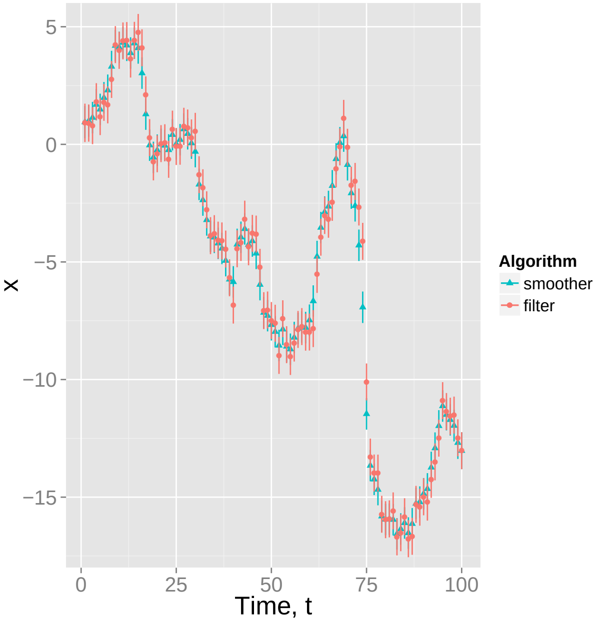
- Univariate Linear Gaussian SSM:

$$\begin{aligned} \text{transition} \quad f(x'|x) =& \mathcal{N}(x'; x, 1) \\ \text{likelihood} \quad g(y|x) =& \mathcal{N}(y; x, 1) \end{aligned}$$

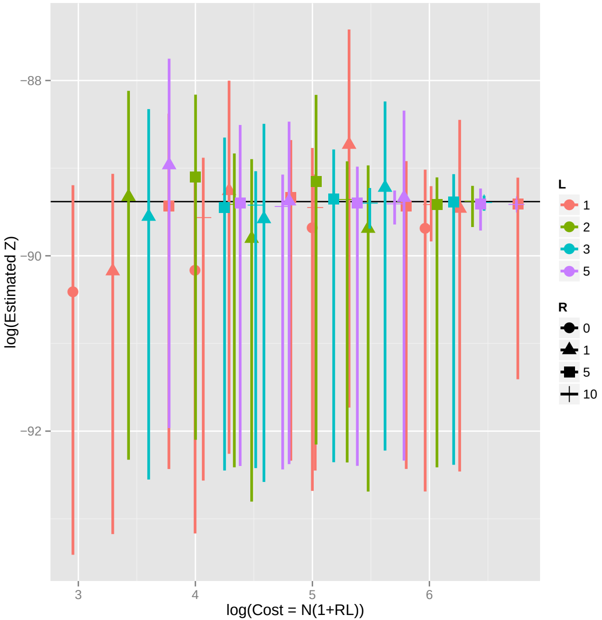
- Artificial jump in observation seqeuence at time 75.
- Cartoon of *model misspecification*
- — a key difficulty with PMCMC.
- Temper only likelihood.
- Use single-site Metropolis-Within Gibbs (standard normal proposal) MCMC moves.



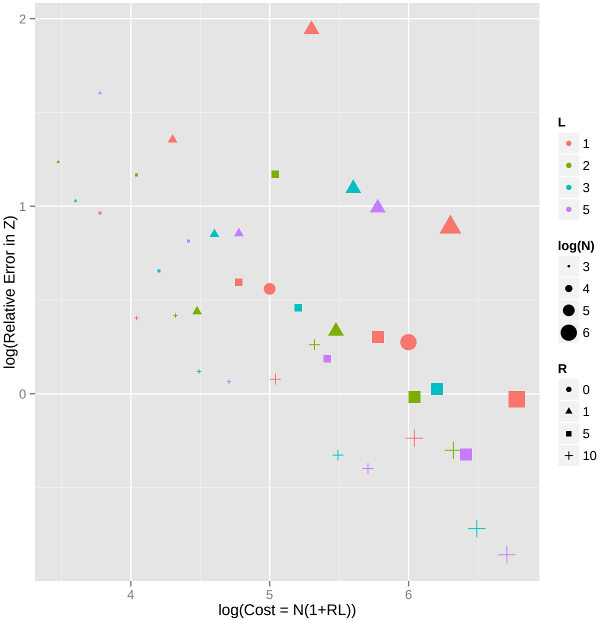
True Filtering and Smoothing Distributions



Estimating the Normalizing Constant, $\widehat{Z}$



Relative Error in $\widehat{Z}$ Against Computational Effort



Performance:

$$(R=1, L=1)$$
$$\prec (R>1, L=1)$$
$$\prec (R>1, L>1)$$

Conclusions

- PMCMC is perhaps even more powerful than has yet been recognised.
- Exploiting the offline nature of the PMCMC setting allows more flexibility than the filtering framework.
- The particular block-tempered approach developed here warrants further investigation.

- Another approach to similar problems is provided by: the iAPF [GJL15, preprint on ArXiv RSN]

Conclusions

- PMCMC is perhaps even more powerful than has yet been widely recognised.
- Exploiting the offline nature of the PMCMC setting allows more flexibility than the filtering framework.
- The particular block-tempered approach developed here warrants further investigation.
- Another approach to similar problems is provided by: the iAPF [GJL15, preprint on ArXiv RSN]

Thanks for listening. . . Any Questions?

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.

C. Andrieu and G. O. Roberts.
The pseudo-marginal approach for efficient Monte Carlo computations.
Annals of Statistics, 37(2):697–725, 2009.

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. Del Moral, A. Doucet, and A. Jasra.
Sequential Monte Carlo samplers.
Journal of the Royal Statistical Society B, 63(3):411–436, 2006.

S. Godsill and T. Clapp.
Improvement strategies for Monte Carlo particle filters.
In A. Doucet, N. de Freitas, and N. Gordon, editors, *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science, pages 139–158. Springer Verlag, New York, 2001.

P. Guarniero, A. M. Johansen, and A. Lee.
The iterated auxiliary particle filter.
In preparation, 2015.

J. Owen, D. J. Wilkinson, and C. S. Gillespie.
Likelihood free inference for markov processes: a comparison.
Statistical Applications in Genetics and Molecular Biology, 2015. In press.