Discretising a Continuous World: Accelerated Inference for State-Space Models via Hidden Markov Models

Mary Llewellyn, Prof. Ruth King,
Dr. Víctor Elvira, Dr. Gordon Ross

THE UNIVERSITY of EDINBURGH

Preprint on arXiv, see [1]
Contact: mary.llewellyn@ed.ac.uk

Main Result

The point mass proposal Metropolis-Hastings (PMPMH) algorithm uses a deterministic grid to induce a hidden Markov model (HMM) approximation to a SSM. The subsequent proposal distribution can be used as an efficient alternative to existing methods that is not susceptible to degeneracy.

State-space models (SSMs)

In discrete time:

(i) an unobserved latent state process, $x_{1:T} = (x_1, \cdots, x_T)$,
(ii) generating time series observations, $y_{1:T} = (y_1, \cdots, y_T)$,
(iii) both endowed with a set of static parameters, $\boldsymbol{\theta}$.

$$\text{State process} \begin{cases} p(x_1 \mid \boldsymbol{\theta}), \\ p(x_t \mid x_{t-1}, \boldsymbol{\theta}), \ t \geq 2, \end{cases}$$

$$\text{Observation process} \left\{ p(y_t \mid x_t, \boldsymbol{\theta}), \ t \geq 1. \right.$$

A **hidden Markov model (HMM)** is a SSM with states in a discrete, finite set.

Computational problem

Data augmentation: specifying the states as additional variables that need to be estimated.

Focus on **Metropolis-Hastings (M-H)** sampling of both $\boldsymbol{\theta}$ and $x_{1:T}$.
(e.g. when $p(y_{1:T} \mid \boldsymbol{\theta})$ and PMMH [2] inefficient, or when we require samples of $(\boldsymbol{\theta}, x_{1:T})$)

Can sample $\boldsymbol{\theta}$ and $x_{1:T}$ from their full conditional distributions (e.g. particle Gibbs [3], also [4]):

1. sample $x_{1:T}^{(i)}$ targeting $p(x_{1:T} \mid y_{1:T}, \boldsymbol{\theta}^{(i-1)})$ (tractable).
2. sample $\boldsymbol{\theta}^{(i)}$ targeting $p(\boldsymbol{\theta} \mid x_{1:T}^{(i)}, y_{1:T})$ (also tractable!).

However, (1) can be challenging to design efficiently if there are many highly correlated states.

Point Mass Proposal Metropolis-Hastings (PMPMH)

A HMM-based M-H proposal distribution.

Step 1: Impose a deterministic grid on the state space and use to approximate the SSM by a discrete HMM with a tractable state-conditional distribution (**Fig. 1**).

Step 2: Propose continuous values for the states using simple bounded distributions (e.g. a uniform distribution). Accept or reject according to their M-H acceptance probability.

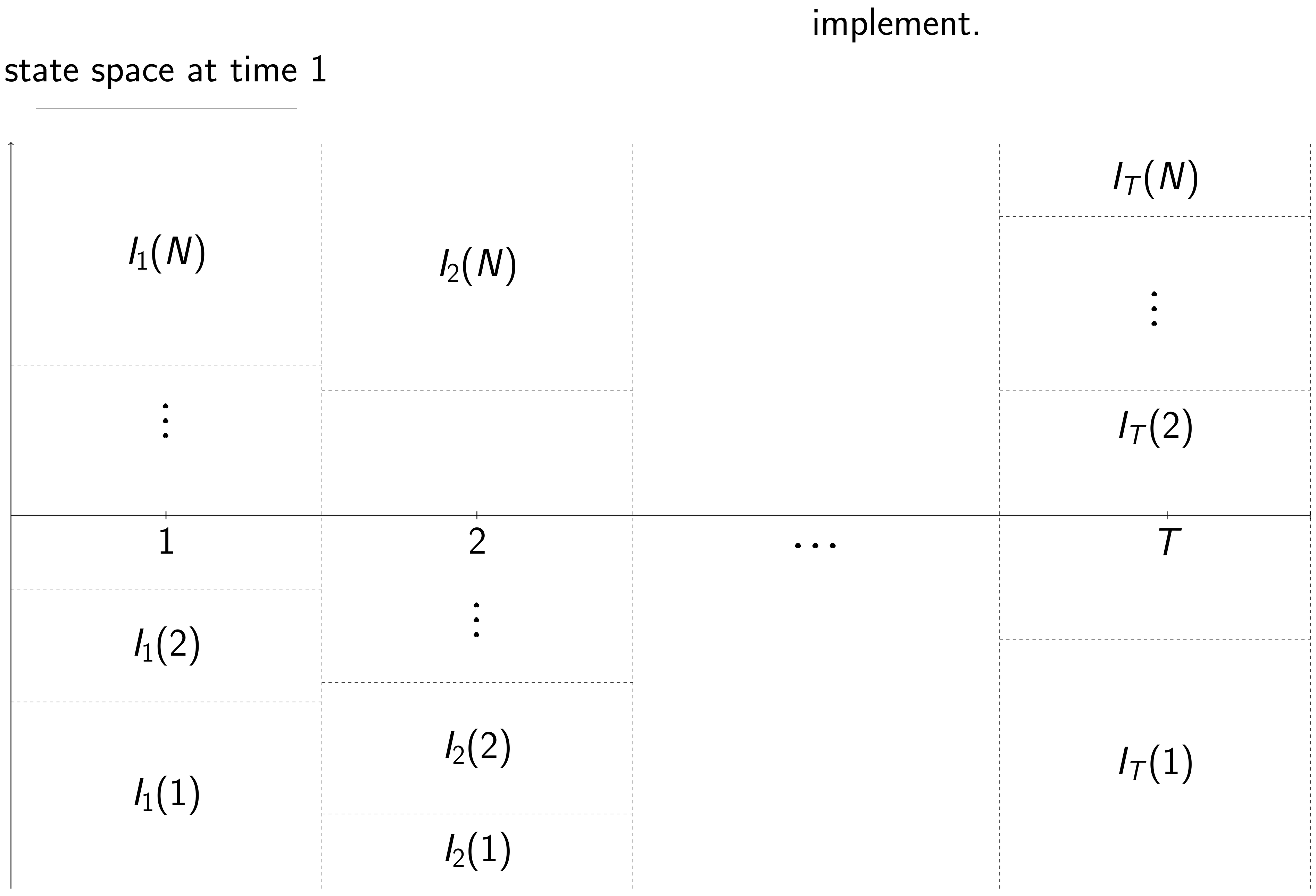

Fig. 1: Example partition of the state space for each time point.

- Focuses the sampler in high density regions but does not degenerate.
- Highly flexible and can be used with efficient block updates of the states.

Implementation

(i) Fewer grid cells often optimal in terms of efficiency (5 – 25).
(ii) Defining the grid at quantiles around the current state usually improves efficiency.
(iii) Efficient blocking strategies easy to implement.

References

1. Llewellyn, M. *et al.* A Point Mass Proposal Method for Bayesian State-Space Model Fitting. *arXiv:2203.13649* (2022).
2. Andrieu, C. *et al.* Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. J Series B.* (2010).
3. Lindsten, F. *et al.* On the use of backward simulation in the particle Gibbs sampler. *2012 IEEE ICASSP* (2012).
4. Frühwirth-Schnatter, S. in *State Space and Unobserved Component Models: Theory and Applications* (2004).
5. Lindsten, F. *et al.* Particle Gibbs with Ancestor Sampling. *J. Mach. Learn. Res.* (2014).
6. Wood, S. N. Statistical inference for noisy nonlinear ecological dynamic systems. *Nature* (2010).

Results: Gaussian mixture model

States follow a Gaussian distribution with high variance (700 units) with low probability. PMPMH provides an efficient approach (**Table 1**) compared to the PGAS algorithm [5], which did not converge with 1000 particles.

Finite-cell span	Number of grid cells	ESS	ESS/s
2 – 4	5	530	0.27
	10	1110	0.28
	20	1030	0.09
4 – 6	5	100	0.06
	10	2490	0.63
	20	3100	0.26

Table 1: Effective sample size (ESS) and ESS/s for a data-centred grid cell PMPMH approach when applied to the Gaussian mixture SSM. Based on 10,000 iterations.

Results: Nicholson's blowfly model

Model for near-chaotic blowfly population growth [6]. Reasonable posterior density estimates for PMPMH vs. PGAS (**Fig. 2**).

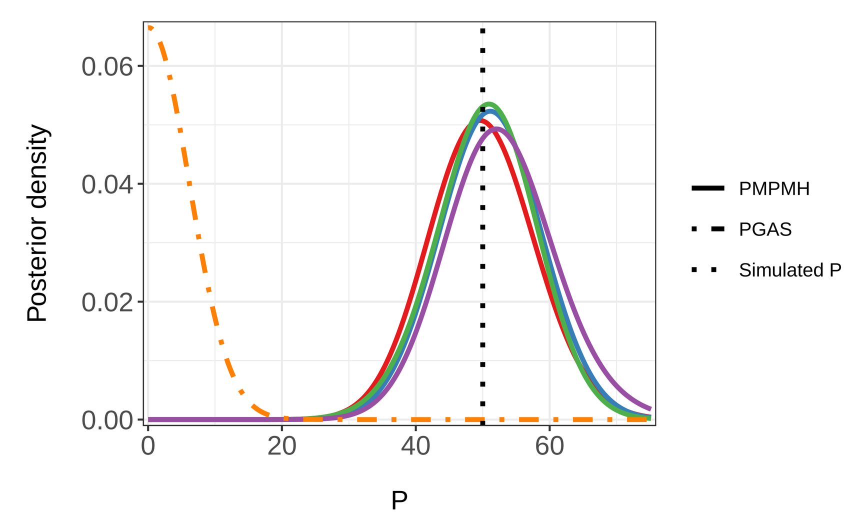

Fig. 2: PMPMH posterior density estimates with 25 grid cells compared to PGAS with 1000 particles (does not converge).