Island filters for inference on metapopulation dynamics, with epidemiological applications

Edward Ionides

University of Michigan, Department of Statistics

Department seminar
Friday March 20, 2020

Background

The curse of dimensionality. Particle filter (PF) methods are effective for inference on low-dimensional nonlinear partially observed stochastic dynamic systems. They scale exponentially badly.

Independent island filters. We study three algorithms that combine many non-interacting Monte Carlo processes.

- Basic island filter (BIF)
- Adapted simulation island filter (ASIF)
- Adapted simulation island filter with intermediate resampling (ASIF-IR)

Metapopulation dynamics. Island filters have theoretical and empirical scaling properties suited to collections of weakly coupled populations.

So, what about COVID-19?

- Researchers have developed many models for disease spread.

- Most of these build on the SIR (Susceptible-Infected-Removed) model that divides a population into three homogeneous classes.

- Extensions can include a latent period after infection, age structure, spatial structure, temperature, control policies.

- We may observe some fraction of cases.

- These models are partially observed stochastic dynamic systems.

- Understanding of COVID-19 epidemiology draws on analysis of previous epidemics combined with assessment of limited available data.

- Methods that fit a general class of mechanistic models assist with formulating and testing scientific hypotheses.

What is a SpatPOMP?

POMP models are partially observed Markov processes, also known as state space models or hidden Markov models
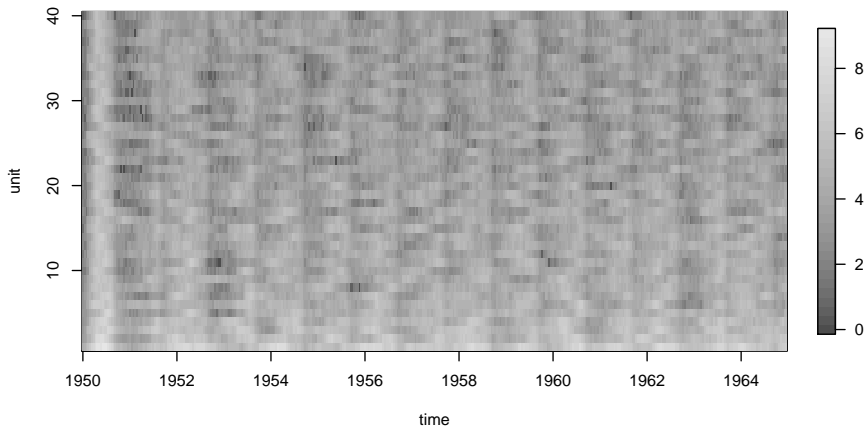
SpatPOMP models are POMP models with a unit structure

Latent Markov process: $X_{u,n} = X_u(t_n)$, $u \in 1 : U$, $n \in 1 : N$

Observation process: $Y_{u,n}$ depends only on $X_{u,n}$

The units could be a metapopulation, say cities in an epidemic model

$U = 40$ units for a coupled measles SEIR model

Simulated data using a gravity model with geography, demography and transmission parameters corresponding to UK pre-vaccination measles.

Island filter inputs, outputs and implicit loops.

input:
simulator for $f_{\boldsymbol{X}_0}(\boldsymbol{x}_0)$ and $f_{\boldsymbol{X}_n|\boldsymbol{X}_{n-1}}(\boldsymbol{x}_n \,|\, \boldsymbol{x}_{n-1})$
evaluator for $f_{Y_{u,n}|X_{u,n}}(y_{u,n} \,|\, x_{u,n})$
number of islands, $\mathcal{I}$
neighborhood structure, $B_{u,n}$
data, $\boldsymbol{y}_{1:N}^*$
ASIF and ASIF-IR: particles per island, J
ASIF-IR: number of intermediate timesteps, S
ASIF-IR: measurement variance parameterizations, $\overset{\leftarrow}{\mathrm{v}}_{u,n}$ and $\overset{\rightarrow}{\mathrm{v}}_{u,n}$
ASIF-IR: approximate process and observation mean functions, $\boldsymbol{\mu}$ and $h_{u,n}$
output:
Log likelihood estimate, $\ell^{\mathsf{MC}} = \sum_{n=1}^{N} \sum_{u=1}^{U} \ell_{u,n}^{\mathsf{MC}}$
implicit loops:
u in $1{:}U$, n in $1{:}N$, i in $1{:}\mathcal{I}$, j in $1{:}J$

BIF. Basic island filter.

Simulate $\boldsymbol{X}_{0:N,i} \sim f_{\boldsymbol{X}_{0:N}}(\boldsymbol{x}_{0:N})$

Measurement weights, $w_{u,n,i}^M = f_{Y_{u,n}|X_{u,n}}(y_{u,n}^* \mid X_{u,n,i})$

Prediction weights, $w_{u,n,i}^P = \prod_{(\tilde{u},\tilde{n}) \in B_{u,n}} w_{\tilde{u},\tilde{n},i}^M$

$\ell_{u,n}^{\mathtt{MC}} = \log \left(\sum_{i=1}^{\mathcal{I}} w_{u,n,i}^M w_{u,n,i}^P \right) - \log \left(\sum_{i=1}^{\mathcal{I}} w_{u,n,i}^P \right)$

The basic island filter is not as naive as it may first appear

- BIF seems naive. Particle filter (PF) method are well known to scale better with N than unconditional simulations.

- BIF scales well with U for weakly coupled systems.

- With modern computers, large numbers of simulations are feasible even when U and N are not small.

- Initially we studied BIF as a theoretical toy. Then we found it is competitive in practice on some models of interest.

Adapted simulation: An easier problem than filtering

- We aim to make each island track the data in a weak sense that does not involve a solution to the full filtering problem.

- The adapted simulation problem is to draw from
  $f_{\boldsymbol{X}_n|\boldsymbol{Y}_n, \boldsymbol{X}_{n-1}}\big(\boldsymbol{x}_n \mid \boldsymbol{y}^*_n, \boldsymbol{x}_{n-1}\big)$.

- The adapted simulation island filter (ASIF) algorithm uses importance sampling to carry out adapted simulation on each island.

- ASIF calculates the likelihood using the proper weight restricted to a neighborhood.

ASIF. Adapted simulation island filter.

Initialize adapted simulation: $\boldsymbol{X}_{0,i}^{\mathrm{A}} \sim f_{\boldsymbol{X}_0}(\boldsymbol{x}_0)$

For n in $1:N$

    Proposals: $\boldsymbol{X}_{n,i,j}^{\mathrm{P}} \sim f_{\boldsymbol{X}_n | \boldsymbol{X}_{n-1}}\big(\boldsymbol{x}_n \mid \boldsymbol{X}_{n-1,i}^{\mathrm{A}}\big)$

    Measurement weights: $w_{u,n,i,j}^M = f_{Y_{u,n} | X_{u,n}}\big(y_{u,n}^* \mid X_{u,n,i,j}^{\mathrm{P}}\big)$

    Adapted resampling weights: $w_{n,i,j}^{\mathrm{A}} = \prod_{u=1}^{U} w_{u,n,i,j}^M$

    Resampling: $\mathbb{P}\big[r(i) = a\big] = w_{n,i,a}^{\mathrm{A}} \Big(\sum_{k=1}^{J} w_{n,i,k}^{\mathrm{A}} \Big)^{-1}$

    $\boldsymbol{X}_{n,i}^{\mathrm{A}} = \boldsymbol{X}_{n,i,r(i)}^{\mathrm{P}}$

    $w_{u,n,i,j}^{\mathrm{P}} = \prod_{\tilde{n}=1}^{n-1} \Big[\frac{1}{J} \sum_{k=1}^{J} \prod_{(\tilde{u},\tilde{n}) \in B_{u,n}^{[\tilde{n}]}} w_{\tilde{u},\tilde{n},i,k}^M \Big] \prod_{(\tilde{u},n) \in B_{u,n}^{[n]}} w_{\tilde{u},n,i,j}^M$

End for

$$\ell_{u,n}^{\mathrm{MC}} = \log\left(\frac{\sum_{i=1}^{\mathcal{I}} \sum_{j=1}^{J} w_{u,n,i,j}^M w_{u,n,i,j}^P}{\sum_{i=1}^{\mathcal{I}} \sum_{j=1}^{J} w_{u,n,i,j}^P} \right)$$

Intermediate resampling

- **Intermediate resampling** splits the time interval between observations into S subintervals.

- Reweighting and/or sampling at each subinterval uses a revised estimate of the anticipated measurement density at the end of the interval called a **guide function**.

- This is applicable to continuous time models.

- Intermediate resampling has useful theoretical and empirical properties (Del Moral and Murray, 2015; Park and Ionides, 2019).

- Intermediate resampling for adapted simulation within ASIF gives the ASIF-IR algorithm.

- Intermediate resampling within PF gives the guided intermediate resampling filter (GIRF) of Park and Ionides (2019), a generalization of the auxiliary particle filter of Pitt and Shepard (1999).

ASIF-IR. ASIF with intermediate resampling.

Initialize adapted simulation: $\boldsymbol{X}_{0,i}^{\mathrm{A}} \sim f_{\boldsymbol{X}_0}(\boldsymbol{x}_0)$

For n in $1\!:\!N$

 Guide simulations: $\boldsymbol{X}_{n,i,j}^{G} \sim f_{\boldsymbol{X}_n | \boldsymbol{X}_{n-1}}\big(\boldsymbol{x}_n \,|\, \boldsymbol{X}_{n-1,i}^{\mathrm{A}}\big)$

 Guide variance: $V_{u,n,i} = \mathrm{Var}\big\{ h_{u,n}\big(X_{u,n,i,j}^{G}\big), j \text{ in } 1\!:\!J \big\}$

 $g_{n,0,i,j}^{\mathrm{R}} = 1$ and $\boldsymbol{X}_{n,0,i,j}^{\mathrm{IR}} = \boldsymbol{X}_{n-1,i}^{\mathrm{A}}$

 For s in $1\!:\!S$

  Intermediate proposals: $\boldsymbol{X}_{n,s,i,j}^{\mathrm{IP}} \sim f_{\boldsymbol{X}_{n,s} | \boldsymbol{X}_{n,s-1}}\big(\,\cdot\,|\,\boldsymbol{X}_{n,s-1,i,j}^{\mathrm{IR}}\big)$

  $\boldsymbol{\mu}_{n,s,i,j}^{\mathrm{IP}} = \boldsymbol{\mu}\big(\boldsymbol{X}_{n,s,i,j}^{\mathrm{IP}}, t_{n,s}, t_n\big)$

  $V_{u,n,s,i,j}^{\mathrm{meas}} = \overrightarrow{\mathrm{v}}_u(\theta, \mu_{u,n,s,i,j}^{\mathrm{IP}})\,,\qquad V_{u,n,s}^{\mathrm{proc}} = V_{u,n,i}\,(t_n - t_{n,s})\big/(t_n - t_{n,0})$

  $\theta_{u,n,s,i,j} = \overleftarrow{\mathrm{v}}_u\big(V_{u,n,s,i,j}^{\mathrm{meas}} + V_{u,n,s,i}^{\mathrm{proc}}, \mu_{u,n,s,i,j}^{\mathrm{IP}}\big)$

  $g_{n,s,i,j} = \prod_{u=1}^{U} f_{Y_{u,n} | X_{u,n}}\big(y_{u,n}^{*} \,|\, \mu_{u,n,s,i,j}^{\mathrm{IP}}; \theta_{u,n,s,i,j}\big)$

  Guide weights: $w_{n,s,i,j}^{G} = g_{n,s,i,j}\big/g_{n,s-1,i,j}^{\mathrm{R}}$

  Resampling: $\mathbb{P}\big[r(i,j) = a\big] = w_{n,s,i,a}^{G}\Big(\sum_{k=1}^{J} w_{n,s,i,k}^{G}\Big)^{-1}$

  $\boldsymbol{X}_{n,s,i,j}^{\mathrm{IR}} = \boldsymbol{X}_{n,s,i,r(i,j)}^{\mathrm{IP}}$ and $g_{n,s,i,j}^{\mathrm{R}} = g_{n,s,i,r(i,j)}$

 End For

 Set $\boldsymbol{X}_{n,i}^{\mathrm{A}} = \boldsymbol{X}_{n,S,i,1}^{\mathrm{IR}}$

 Measurement weights: $w_{u,n,i,j}^{M} = f_{Y_{u,n} | X_{u,n}}\big(y_{u,n}^{*} \,|\, X_{u,n,i,j}^{G}\big)$

 $w_{u,n,i,j}^{\mathrm{P}} = \prod_{\tilde{n}=1}^{n-1}\Big[\frac{1}{J}\sum_{a=1}^{J}\prod_{(\tilde{u},\tilde{n}) \in B_{u,n}^{[\tilde{n}]}} w_{\tilde{u},\tilde{n},i,a}^{M}\Big]\prod_{(\tilde{u},n) \in B_{u,n}^{[n]}} w_{\tilde{u},n,i,j}^{M}$

End for

$\ell_{u,n}^{\mathrm{MC}} = \log\left(\dfrac{\sum_{i=1}^{\mathcal{I}}\sum_{j=1}^{J} w_{u,n,i,j}^{M} w_{u,n,i,j}^{\mathrm{P}}}{\sum_{i=1}^{\mathcal{I}}\sum_{j=1}^{J} w_{u,n,i,j}^{\mathrm{P}}}\right)$


Software for SpatPOMP models

- We use the `asif`, `asifir` and `girf` implementations in the R package `spatPomp` (Asfaw et al., 2019).

- `spatPomp` offers a class 'spatPomp' that extends the 'pomp' class for POMP models in the R package pomp (King et al., 2016).

- All methods available in pomp can formally be applied to 'spatPomp' objects, though they may not be practically effective for spatiotemporal POMPs.

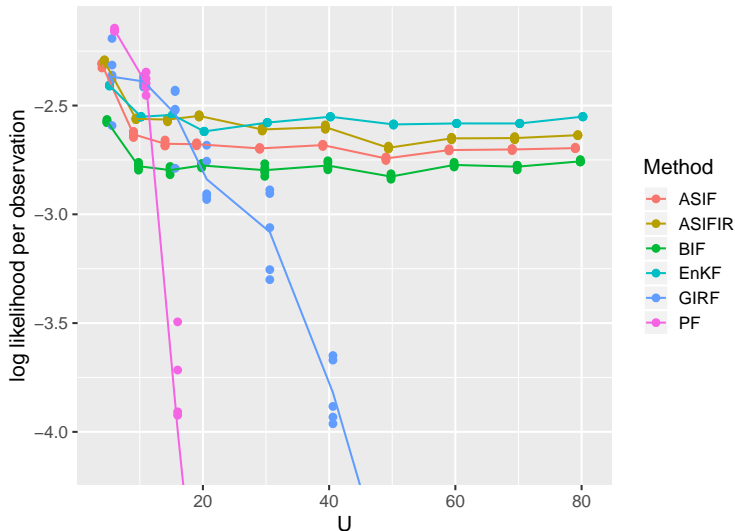# Filtering U-dimensional correlated Brownian motion
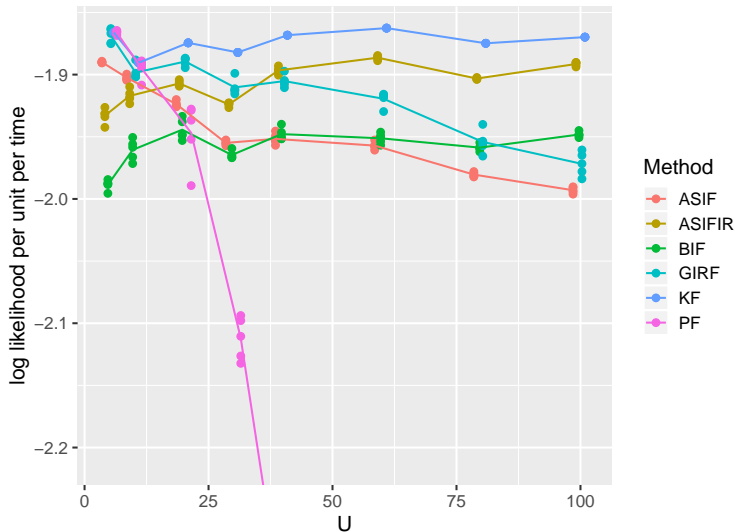
$$\text{Cov}\big(X_{u,n} - X_{u,n-1}, X_{\tilde{u},n} - X_{\tilde{u},n-1}\big) \sim 0.4^{|u-\tilde{u}|}$$

Filtering U units of a coupled measles SEIR model
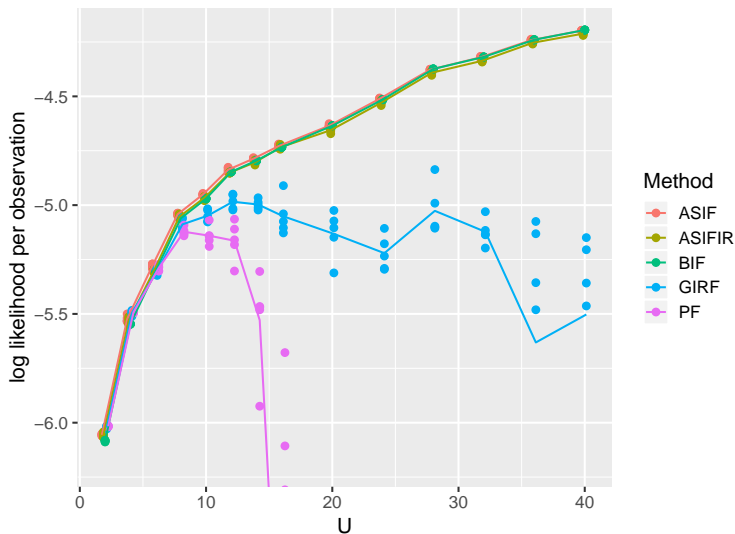
Simulated data using a gravity model with geography, demography and transmssion parameters corresponding to UK pre-vaccination measles.

# Filtering U units of Lorenz 96 toy atmospheric model

$$dX_u(t) = \left\{ X_{u-1}(t)\big(X_{u+1}(t) - X_{u-2}(t)\big) - X_u(t) + F \right\} dt + \sigma\, dB_u(t)$$

From filtering to parameter inference

- Log likelihood evaluation in principle enables likelihood-based or Bayesian inference.

- Iterated filtering maximizes the likelihood for PF or GIRF (Ionides et al., 2015).

- Particle Markov chain Monte Carlo can be applied with any likelihood estimate (Andrieu et al., 2010). It is numerically intractable when Monte Carlo estimates are costly and noisy.

- Extending iterated filtering to island filters is future work.

Measles likelihood slices for G and μ_{IR} via ASIF
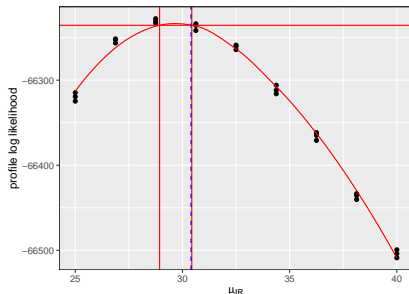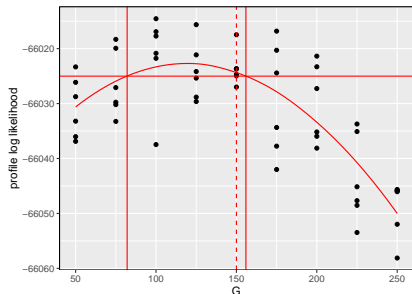
- Simulating 15 year of data from $U = 40$ cities for the measles model.

- The gravitational coupling constant G is fairly weakly identified: a week of computing on a 30 core machine gives Monte Carlo error on the same scale as the statistical uncertainty.

- The recovery rate μ_{IR} is well identified.

Theorem

Let ℓ^{MC} denote the Monte Carlo likelihood approximation constructed by BIF, ASIF or ASIF-IR. Consider a limit with a growing number of islands, $\mathcal{I} \to \infty$. Suppose regularity assumptions listed in the arXiv preprint. There are quantities $\epsilon(U, N) = O(1)$ and $V(U, N) = O(U^2 N^2)$ such that

$$\mathcal{I}^{1/2} \big[\ell^{MC} - \ell - \epsilon U N \big] \xrightarrow[\mathcal{I} \to \infty]{d} \mathcal{N} \big[0, V \big],$$

where $\xrightarrow[\mathcal{I} \to \infty]{d}$ denotes convergence in distribution and $\mathcal{N}[\mu, \Sigma]$ is the normal distribution with mean μ and variance Σ. If an additional spatiotemporal mixing assumption holds, we obtain an improved variance bound

$$V(U, N) = O(U N)$$

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

Asfaw, K., Ionides, E. L., and King, A. A. (2019). `spatPomp`: R package for statistical inference for spatiotemporal partially observed Markov processes. *https: // github. com/ kidusasfaw/ spatPomp* .

Del Moral, P. and Murray, L. M. (2015). Sequential Monte Carlo with highly informative observations. *Journal on Uncertainty Quantification*, 3:969–997.

Ionides, E. L., Nguyen, D., Atchadé, Y., Stoev, S., and King, A. A. (2015). Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proceedings of the National Academy of Sciences of the USA*, 112:719–724.

King, A. A., Nguyen, D., and Ionides, E. L. (2016). Statistical inference for partially observed Markov processes via the R package pomp. *Journal of Statistical Software*, 69:1–43.

Park, J. and Ionides, E. L. (2019). Inference on high-dimensional implicit dynamic models using a guided intermediate resampling filter. *Arxiv:1708.08543v3*.

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