

# **An iterated block particle filter for inference on coupled dynamic systems**

Edward Ionides

University of Michigan, Department of Statistics

IMS Asia Pacific Rim Meeting

Jan 5, 2024

Slides are at <https://ionides.github.io/talks/aprm24.pdf>

Joint work with Patricia Ning, Jesse Wheeler, Kidus Asfaw, Jifan Li,  
Joonha Park, Aaron King, Mercedes Pascual

# Question to be addressed

- 1 When can we carry out full-information likelihood-based inference on nonlinear non-Gaussian spatiotemporal partially observed Markov process (SpatPOMP) models? In particular, models for networks of interacting biological population dynamics.
- 2 We introduce the **iterated block particle filter**, currently the most powerful algorithm available in the spatPomp R package.

# Inference challenges in population dynamics

- ① Combining measurement noise and process noise.
- ② Including covariates in mechanistically plausible ways.
- ③ Continuous time models.
- ④ Modeling and estimating interactions in coupled systems.
- ⑤ Dealing with unobserved variables.
- ⑥ **Modeling spatiotemporal dynamics.**
- ⑦ Studying population dynamics via genetic sequence data.

1–5 are largely solved, from a methodological perspective.

6 is our immediate topic.

7 is exciting but not the focus of this talk.

Reviews: Bjornstad & Grenfell (*Science*, 2001); Grenfell et al (*Science*, 2004)

- Consideration of arbitrary dynamic models. The limitations should be our scientific creativity and the information in the data.

Hence, **plug-and-play** methods which need a simulator from the model but not nice closed-form expressions for densities.

- Statistically efficient inference, to extract all the information in the data.

Hence, **likelihood-based** methods.

# Fitting mechanistic models to time series

- Iterated particle filtering via `mif2` in the R package `pomp` enables Masters-level statisticians to do plug-and-play likelihood-based inference:

<https://ionides.github.io/531w22/>

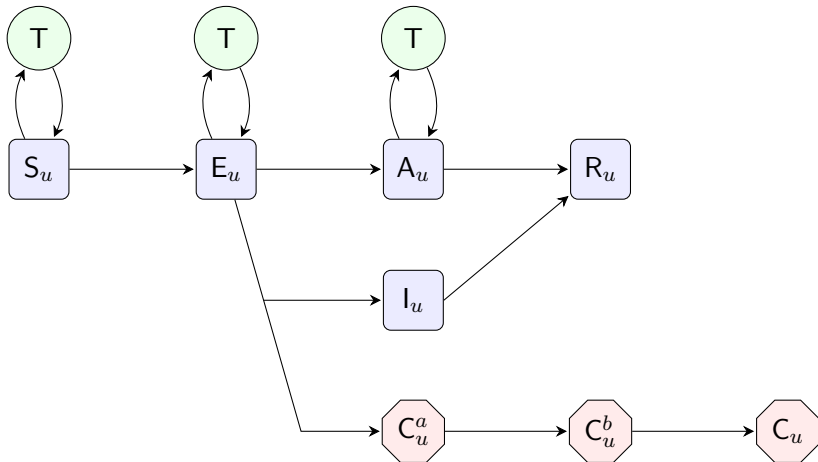
- The science may be hard, but the statistics is becoming routine.

# The curse of dimensionality

- Particle filter (PF) methods fail for high-dimensional systems. They scale exponentially badly.
- Algorithms with improved scalability include:
  - Bagged filters (BF, IBF)**
  - Ensemble Kalman filter (EnKF, IEnKF)**
  - Guided intermediate resampling filter (GIRF, IGIRF)**
  - Block particle filter (BPF, IBPF)**
- Filters estimate latent states and evaluate the likelihood.
- Iterated filters estimate parameters using stochastic parameter perturbations.
- These algorithms are all implemented in the `spatPomp` R package.

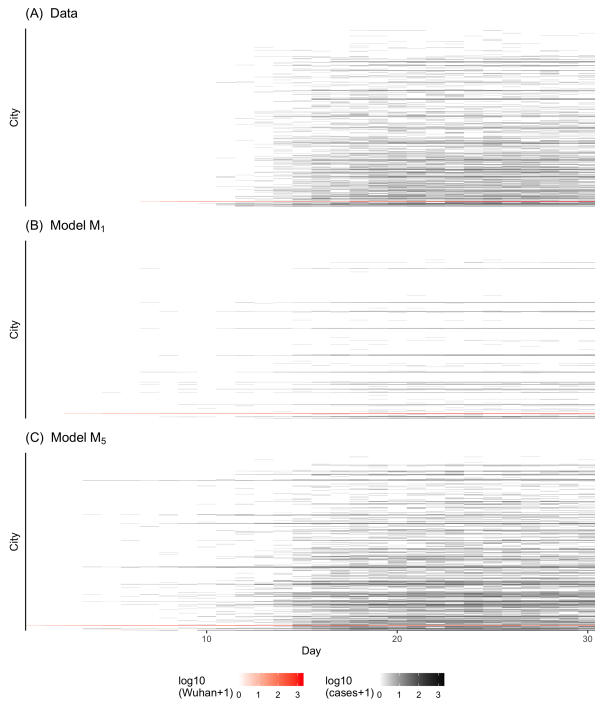
# COVID in 375 cities in China, 10-23 January, 2020

- Metapopulation data were used to infer the fraction of asymptomatic cases and their contagiousness (Li et al, *Science*, May 2020).
- SEIR (susceptible-exposed-infected-removed) model with asymptomatics, reporting delay, and coupling based on cell phone data.
- Li et al (2020) used iterated EnKF for inference.
- The time interval covers the initial China lockdown.
- Here, we summarize our re-analysis of this model and data, recently posted on arXiv.



- Reportably infectious individuals,  $I_u$  for city  $u$ , are included in the delayed reporting compartment,  $C_u^a$ .
- An individual arriving at  $C_u$  is a case report for city  $u$ .
- Individuals in  $A_u$  are not reportable and transmit at a reduced rate.
- Travel occurs to and from  $T$ , based on 2018 data from Tencent.





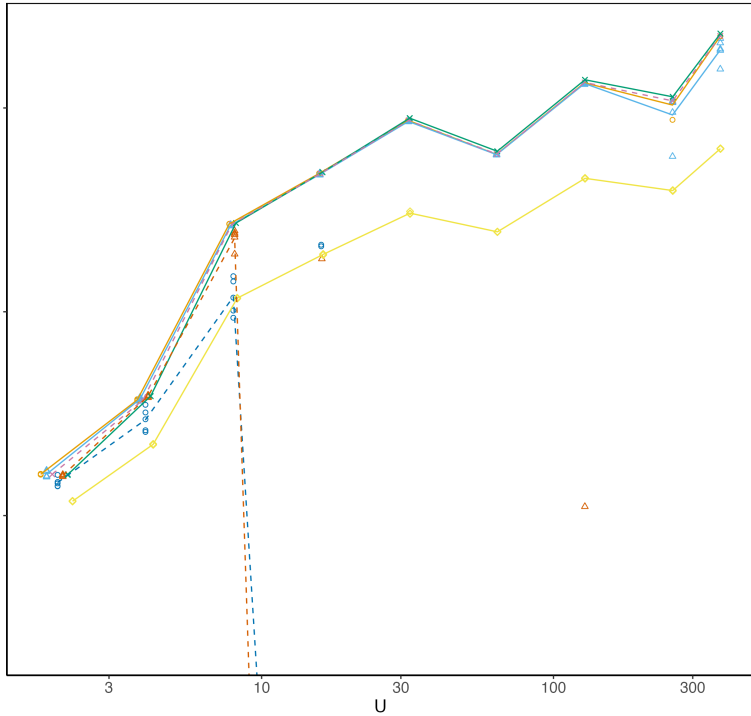
Log-likelihood per unit per time

U

Method

- ABF
- ABF-IR
- BPF
- EnKF
- GIRF
- PF
- UBF

10



# More on the block particle filter (BPF)

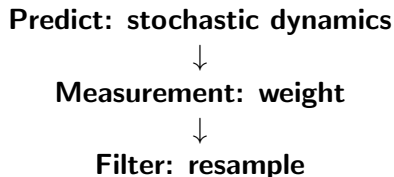
- BPF also worked quickly, easily and reliably on a measles metapopulation, as well as various toy benchmark problems (Ionides et al, *JASA*, 2023).
- BPF has theoretical support in some situations (Rebeschini & Van Handel, *Annals of Applied Probability*, 2015).
- This motivated us to develop an IBPF for parameter estimation.
- IBPF has theoretical guarantees similar to BPF (Ning & Ionides, *JMLR*, 2023).
- BPF was independently proposed as the “factored particle filter” by Ng et al (2002).

# Particle filter (PF)

## Evolutionary analogy



## Particle filter algorithm

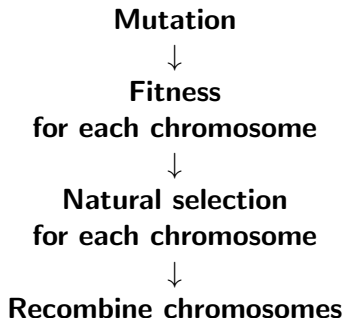


- PF is an evolutionary algorithm with good mathematical properties: an unbiased likelihood estimate and consistent latent state distribution.

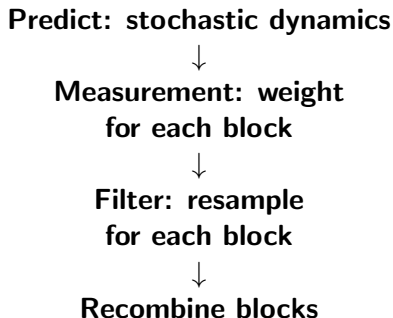
# Block particle filter (BPF)

- Blocks are a partition of the metapopulation units.
- For measles, we use each city as a block.

## Evolutionary analogy



## Block particle filter

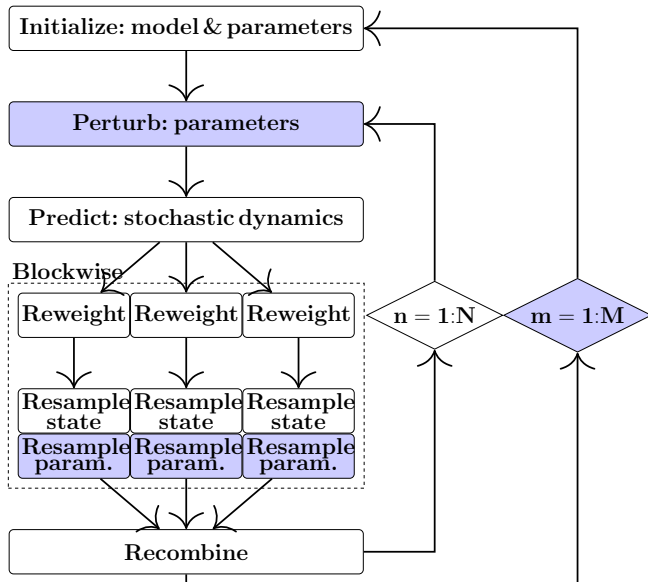


- Blocks are segments of the full state which can be reassorted between particles at the resampling step.

# Comments on the Ensemble Kalman Filter (EnKF)

- EnKF is more dependent on approximate Gaussianity than is sometimes supposed.
- The Gaussian-inspired update rule is similar to the extended Kalman filter (EKF), which has largely been superseded by particle filter methods for low-dimensional nonlinear biological dynamics.
- Simple systems can defeat EnKF: the linear Gaussian update is helpless when data inform the conditional variance rather than the conditional mean.
- Big systems need computationally tractable analysis. EnKF may sometimes be the best solution available, but be aware of its limitations.

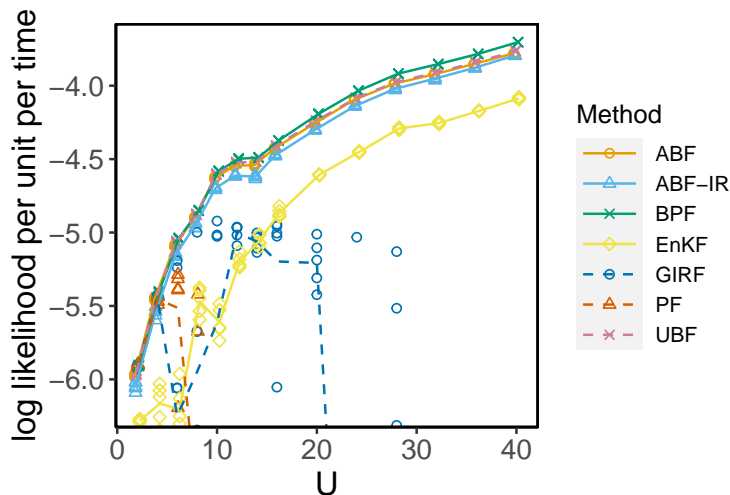
# An iterated block particle filter for parameter estimation



- ① Monte Carlo adjusted profile likelihood (?) obtains confidence intervals that accommodate Monte Carlo error.
- ② Comparing the log-likelihood with an autoregressive model (or other simple statistical model) provides a check of model fit.
- ③ Comparing the block log-likelihood against the benchmark provides insight into problematic units.
- ④ Comparing the conditional log-likelihood for each observation against the benchmark helps to identify outliers.

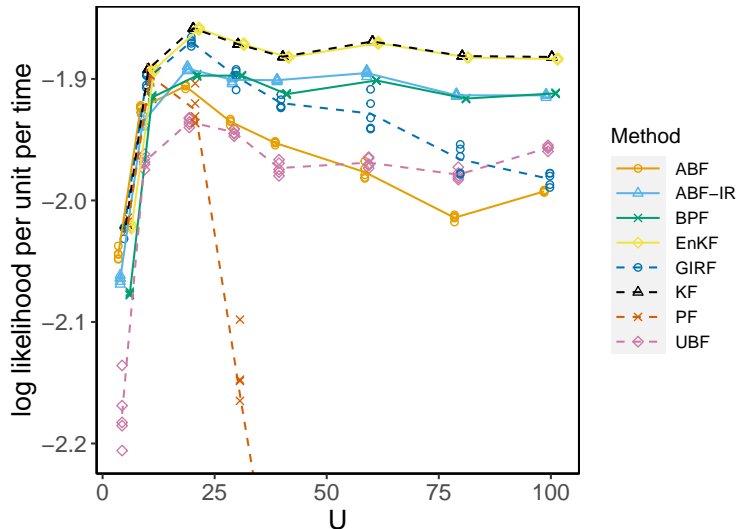


# Filtering $U$ units of a coupled measles SEIR model



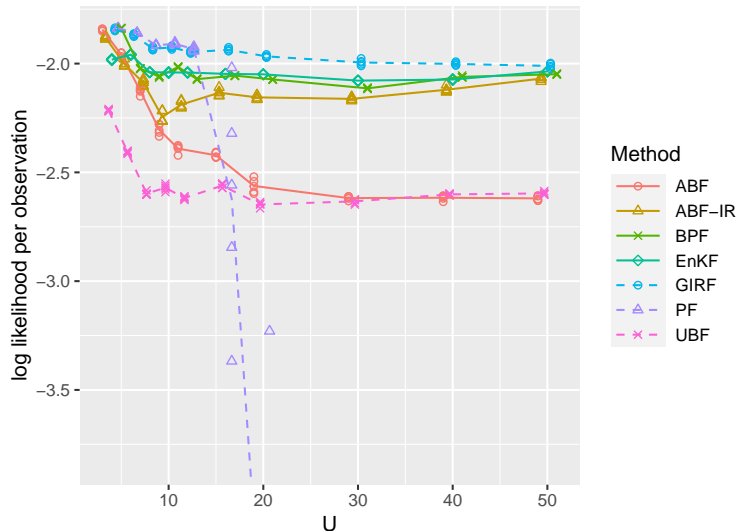
Simulated data using a gravity model with geography, demography and transmission parameters corresponding to UK pre-vaccination measles (Ionides et al, JASA, 2023).

# Filtering $U$ -dimensional correlated Brownian motion



$$\text{Cov}(X_{u,n} - X_{u,n-1}, X_{\tilde{u},n} - X_{\tilde{u},n-1}) \sim 0.4^{|u-\tilde{u}|}$$

# Filtering $U$ units of Lorenz 96 toy atmospheric model



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

# References I

- Asfaw, K., Park, J., King, A. A., and Ionides, E. L. (2023). Partially observed Markov processes with spatial structure via the R package spatPomp. *arXiv:2101.01157v3*.
- Bjørnstad, O. N. and Grenfell, B. T. (2001). Noisy clockwork: Time series analysis of population fluctuations in animals. *Science*, 293:638–643.
- Bretó, C., Ionides, E. L., and King, A. A. (2019). Panel data analysis via mechanistic models. *Journal of the American Statistical Association*, 115:1178–1188.
- Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. A., and Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*, 303:327–332.
- He, D., Ionides, E. L., and King, A. A. (2010). Plug-and-play inference for disease dynamics: Measles in large and small towns as a case study. *Journal of the Royal Society Interface*, 7:271–283.

## References II

- Ionides, E. L., Asfaw, K., Park, J., and King, A. A. (2021). Bagged filters for partially observed interacting systems. *Journal of the American Statistical Association*, pre-published online.
- Ionides, E. L., Ning, N., and Wheeler, J. (2022). An iterated block particle filter for inference on coupled dynamic systems with shared and unit-specific parameters. *Statistica Sinica*, pre-published online.
- Katzfuss, M., Stroud, J. R., and Wikle, C. K. (2020). Ensemble Kalman methods for high-dimensional hierarchical dynamic space-time models. *Journal of the American Statistical Association*, 115(530):866–885.
- Lee, E. C., Chao, D. L., Lemaitre, J. C., Matrajt, L., Pasetto, D., Perez-Saez, J., Finger, F., Rinaldo, A., Sugimoto, J. D., Halloran, M. E., Longini, I. M., Ternier, R., Vissieres, K., Azman, A. S., Lessler, J., and Ivers, L. C. (2020). Achieving coordinated national immunity and cholera elimination in Haiti through vaccination: A modelling study. *The Lancet Global Health*, 8(8):e1081–e1089.

## References III

- Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., and Shaman, J. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, 368(6490):489–493.
- Ng, B., Peshkin, L., and Pfeffer, A. (2002). Factored particles for scalable monitoring. *Proceedings of the 18th Conference on Uncertainty and Artificial Intelligence*, pages 370–377.
- Ning, N. and Ionides, E. L. (2023). Iterated block particle filter for high-dimensional parameter learning: Beating the curse of dimensionality. *Journal of Machine Learning Research*, to appear.
- Park, J. and Ionides, E. L. (2020). Inference on high-dimensional implicit dynamic models using a guided intermediate resampling filter. *Statistics and Computing*, 30(5):1497–1522.

- Rebeschini, P. and van Handel, R. (2015). Can local particle filters beat the curse of dimensionality? *The Annals of Applied Probability*, 25:2809–2866.
- Wheeler, J., Rosengart, A. L., Jiang, Z., Tan, K., Treutle, N., and Ionides, E. L. (2023). Informing policy via dynamic models: Cholera in Haiti. *arXiv:2301.08979*.