Exchangeable Particle Gibbs for Markov Jump Processes

Lanya Yang
Supervisors: Lloyd Chapman and Chris Sherlock

Nov 28, 2025

STOR-i | Lancaster University

Reaction networks

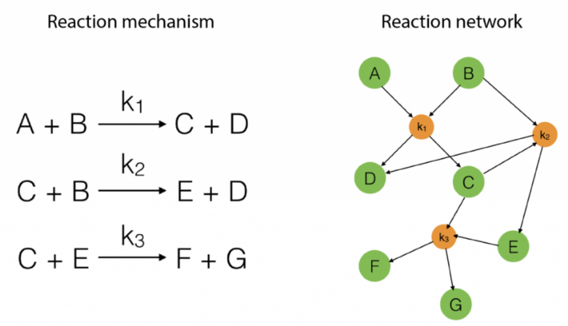


Figure 1: An example of a simple reaction mechanism

Reaction networks

$$\text{Susceptible} \xrightarrow{\beta SI} \text{Infectious} \xrightarrow{\gamma I} \text{Recovered}$$

$$\text{Reaction 1 } (R_1) : S + I \xrightarrow{\beta SI} 2I \quad \text{(Infection)},$$
$$\text{Reaction 2 } (R_2) : I \xrightarrow{\gamma I} R \quad \text{(Recovery)},$$

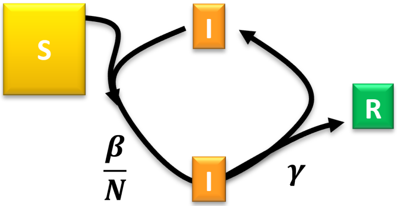


Figure 2: SIR model

Reaction networks

Setup:

- u species: $\mathcal{X}_1, \ldots, \mathcal{X}_u$
- v reactions: $\mathcal{R}_1, \ldots, \mathcal{R}_v$

General form of reaction $\mathcal{R}_i$:

$$\sum_{j=1}^{u} a_{ij}\mathcal{X}_j \xrightarrow{h_i} \sum_{j=1}^{u} b_{ij}\mathcal{X}_j$$

Markov jump process

- Describes how a reaction network evolves over time
- A **continuous-time**, **discrete-state** stochastic process
- Each jump corresponds to the occurrence of a reaction

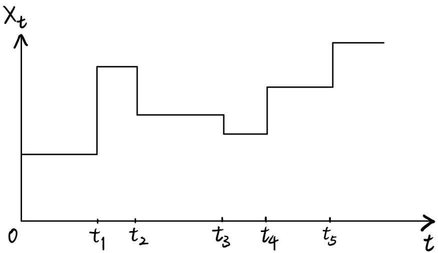


Figure 3: A Markov jump process

Hidden Markov Model (HMM)

$$
\begin{array}{ccccccccc}
X_0 & \xrightarrow{p} & X_{t_1} & \xrightarrow{p} & X_{t_2} & \xrightarrow{p} & \dots & \xrightarrow{p} & X_{t_L} \\
\downarrow g & & \downarrow g & & \downarrow g & & & & \downarrow g \\
Y_0 & & Y_1 & & Y_2 & & & & Y_L
\end{array}
$$

Figure 4: A hidden Markov model with states $X_{t_{\mathbf{0}:L}}$ and observations $Y_{0:L}$

Hidden Markov Model (HMM)

$$\begin{array}{ccccccccc}
X_0 & \xrightarrow{p} & X_{t_1} & \xrightarrow{p} & X_{t_2} & \xrightarrow{p} & \dots & \xrightarrow{p} & X_{t_L} \\
\downarrow g & & \downarrow g & & \downarrow g & & & & \downarrow g \\
Y_0 & & Y_1 & & Y_2 & & & & Y_L
\end{array}$$

Figure 4: A hidden Markov model with states $X_{t_{0:L}}$ and observations $Y_{0:L}$

The mathematical form of the Hidden Markov model is given by

$$\begin{aligned}
X_0 &\sim p_0(\cdot \mid \theta), \\
X_{t_\ell} \mid (x_{[0,t_{\ell-1}]}, y_{0:\ell-1}, \theta) &\sim p(\cdot \mid x_{t_{\ell-1}}, \theta), \quad \ell = 1, \dots, L \\
Y_\ell \mid (x_{[0,t_\ell]}, y_{0:\ell-1}, \theta) &\sim g(\cdot \mid x_{t_\ell}, \theta).
\end{aligned}$$

Bayesian Inference

Given:

- A sequence of observations: $(y_0, y_1, \ldots, y_L)$

Goal:

- Estimate the model parameter, θ
- Estimate the latent states $x_{t_0}, \ldots, x_{t_L}$
- Target distribution:

$$p(x_{t_{0:L}}, \theta \mid y_{0:L})$$

Inference methods

- Approximate Bayesian Computation (ABC)
- Traditional MCMC methods
- Particle MCMC (Andrieu et al.; 2010)
  - Particle Marginal Metropolis-Hastings (PMMH)
  - Particle Gibbs

Particle Filter

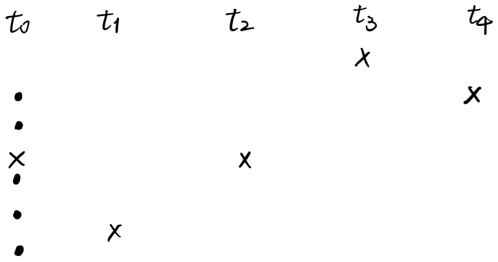


Figure 5: particle filter

The number of proposed particles at each observation time point is denoted by M. Here $M = 5$

Particle Filter

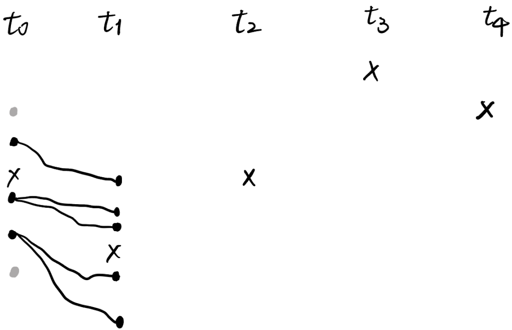


Figure 6: particle filter

Particle Filter

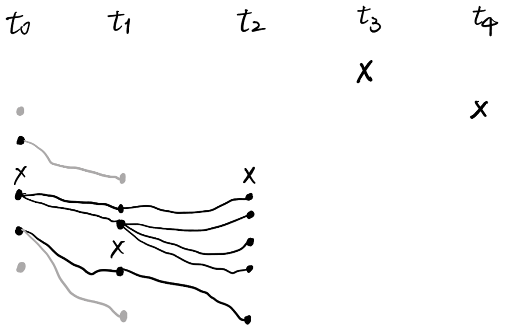


Figure 7: particle filter

Particle Filter

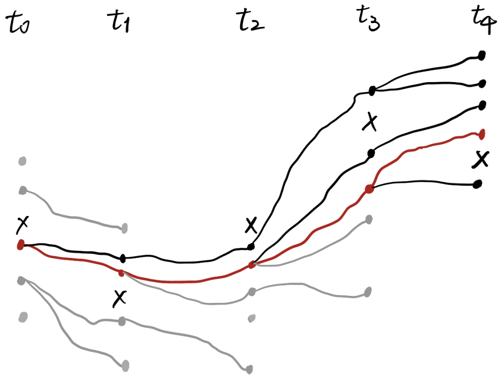


Figure 8: particle filter

Conditional Particle filter

Suppose the number of proposed paths is $M = 4$,

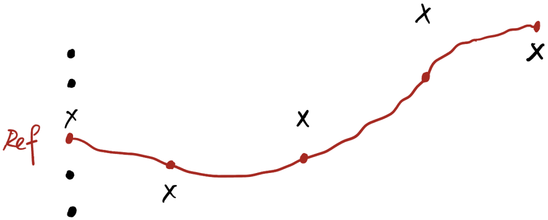


Figure 9: conditional particle filter

Conditional Particle filter

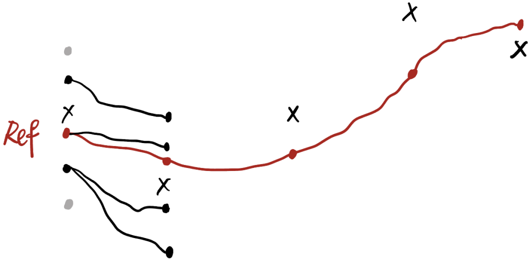


Figure 10: conditional particle filter

Conditional Particle filter

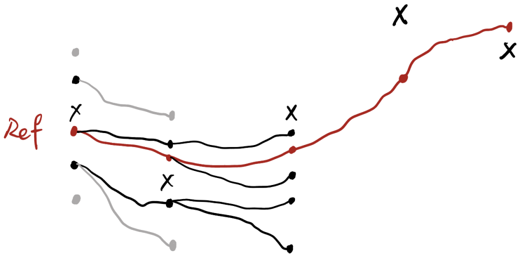


Figure 11: conditional particle filter

Conditional Particle filter

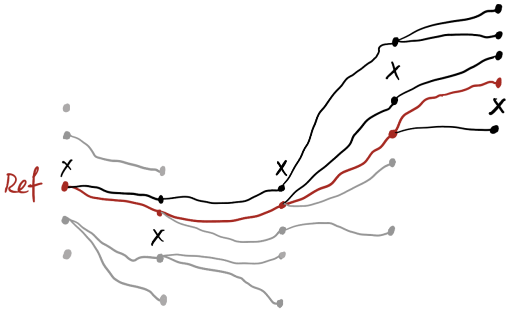


Figure 12: conditional particle filter

Particle Gibbs (PG)

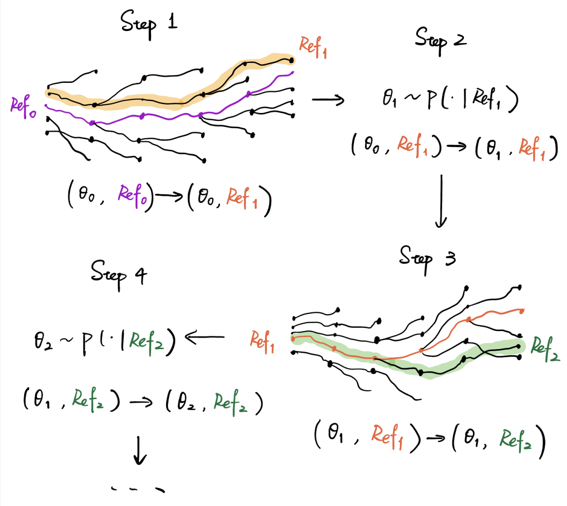


Figure 13: Particle Gibbs sampler

Particle degeneracy

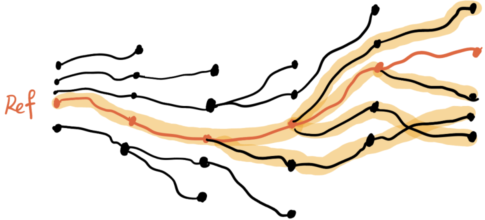


Figure 14: Particle degeneracy

Particle degeneracy

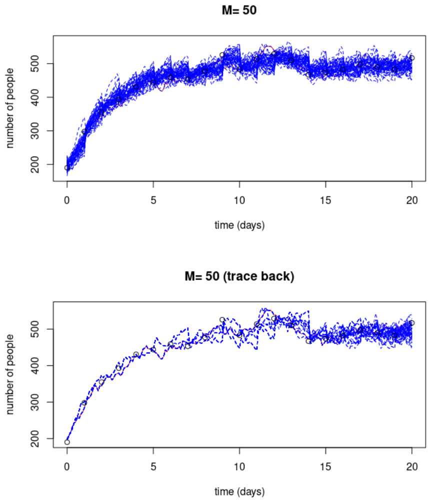
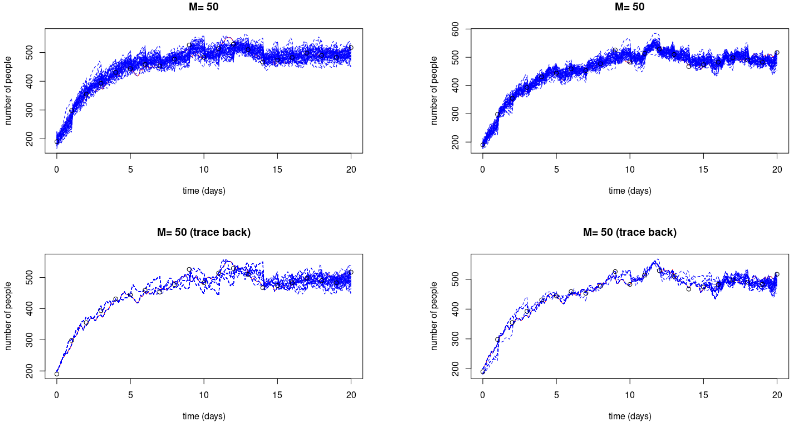


Figure 15: A run of conPF on the SIS model with $\lambda = 1.2$ and $\mu = 0.6$

Methods for addressing particle degeneracy issue

- Particle Gibbs with Ancestor Sampling (PGAS) (Lindsten et al.; 2014)
- Exchangeable Particle Gibbs (xPG)(Malory; 2021)
- Particle-RWM (pRWM)(Finke and Thiery; 2023)
- Particle-MALA (pMALA) (Corenflos and Finke; 2024)

Exchangeable conditional particle filter



conPF run | Exchangeable conPF run

Figure 16: Comparison of conPF and exchangeable conPF on the SIS model with $\lambda = 1.2$ and $\mu = 0.6$.

Exchangeable Particle Gibbs (xPG)

Algorithm 1 Tau-leap method

Choose a step size τ and set initial state x_0 at $t = 0$
while $t < T$ **do**
    Compute reaction hazards $h_i(x_t),\ i = 1, \ldots, v$
    Sample reaction counts $N^{\mathcal{R}_i} \sim \text{Poisson}(h_i(x_t)\tau)$
    Update state: $x_{t+\tau} = x_t + \sum_{i=1}^{v} N^{\mathcal{R}_i} S^i$
    Advance time: $t \leftarrow t + \tau$
end while

Notations:

- $X_t^{(m)}$: state of the m-th proposed path at time t
- $N_k^{(m)}$: number of reactions in the m-th proposed path in the k-th time step

Notations:

- $X_t^{(m)}$: state of the m-th proposed path at time t
- $N_k^{(m)}$: number of reactions in the m-th proposed path in the k-th time step

Step 1. Sample the initial states for the proposed paths $x_0^{(1:M)}$ jointly from $\tilde{q}(\cdot \mid x_0^{(0)})$, such that

$$p_0(x_0^{(0)}|\theta)\tilde{q}_0(x_0^{(1:M)}|x_0^{(0)},\theta) = p_0(x_0^{(j)}|\theta)\tilde{q}_0(x_0^{(-j)}|x_0^{(j)},\theta), \quad \forall j \in \{1,\ldots,M\}, \tag{1}$$

Notations:

- $X_t^{(m)}$: state of the m-th proposed path at time t
- $N_k^{(m)}$: number of reactions in the m-th proposed path in the k-th time step

Step 1. Sample the initial states for the proposed paths $x_0^{(1:M)}$ jointly from $\tilde{q}(\cdot \mid x_0^{(0)})$, such that

$$p_0(x_0^{(0)}|\theta)\tilde{q}_0(x_0^{(1:M)}|x_0^{(0)},\theta) = p_0(x_0^{(j)}|\theta)\tilde{q}_0(x_0^{(-j)}|x_0^{(j)},\theta), \quad \forall j \in \{1,\ldots,M\}, \tag{1}$$

Step 2. For each $k = 1, 2, \ldots$:

Simulate the number of reaction events of the k-th time step, i.e., $N_k^{(1:M)}$ jointly given the number of such events in the reference path $N_k^{(0)}$ and Poisson means $\mu_k^{(0:M)}$.

$$p\left(N_k^{(0)}|\mu_k^{(0)}\right) q\left(N_k^{(1:M)} \mid \mu_k^{(0:M)}, N_k^{(0)}\right) = p\left(N_k^{(j)}|\mu_k^{(j)}\right) q\left(N_k^{(-j)} \mid \mu_k^{(0:M)}, N_k^{(j)}\right). \tag{2}$$

Exchangeable Particle Gibbs (xPG)

Given $N^{(0)} \sim \text{Pois}(\lambda^{(0)})$, we want to construct $N^{(1)} \sim \text{Pois}(\lambda^{(1)})$ such that $N^{(0)}$ and $N^{(1)}$ are correlated.

Case 1: $\lambda^{(1)} > \lambda^{(0)}$ (Poisson update)

$$N^{(1)} = N^{(0)} + \text{Pois}(\lambda^{(1)} - \lambda^{(0)})$$

Exchangeable Particle Gibbs (xPG)

Given $N^{(0)} \sim \text{Pois}(\lambda^{(0)})$, we want to construct $N^{(1)} \sim \text{Pois}(\lambda^{(1)})$ such that $N^{(0)}$ and $N^{(1)}$ are correlated.

Case 1: $\lambda^{(1)} > \lambda^{(0)}$ (Poisson update)

$$N^{(1)} = N^{(0)} + \text{Pois}(\lambda^{(1)} - \lambda^{(0)})$$

Case 2: $\lambda^{(1)} < \lambda^{(0)}$ (Binomial thinning)

$$N^{(1)} \sim \text{Bin}\left(N^{(0)}, \frac{\lambda^{(1)}}{\lambda^{(0)}}\right)$$

xPG

Suppose the number of proposed paths is $M = 4$. On the time interval $[(k-1)\tau, k\tau]$, given $X_{(k-1)\tau}^{(0:4)}$ and θ, we compute the Poisson means of all paths and order them as:

$$\mu^{(2)} < \mu^{(3)} < \mu^{(0)} < \mu^{(4)} < \mu^{(1)}$$

where $\mu^{(0)}$ is the Poisson mean of the reference path.

xPG

Suppose the number of proposed paths is $M = 4$. On the time interval $[(k-1)\tau, k\tau]$, given $X^{(0:4)}_{(k-1)\tau}$ and θ, we compute the Poisson means of all paths and order them as:

$$\mu^{(2)} < \mu^{(3)} < \mu^{(0)} < \mu^{(4)} < \mu^{(1)}$$

where $\mu^{(0)}$ is the Poisson mean of the reference path.

For each reaction type, given the number of such events in the reference path, $N^{(0)}$, simulate the corresponding number of events in the proposed paths.

$$N^{(2)} \leftarrow N^{(3)} \leftarrow N^{(0)} \xrightarrow[+\text{Pois}(\mu^{(4)}-\mu^{(0)})]{} N^{(4)} \xrightarrow[+\text{Pois}(\mu^{(1)}-\mu^{(4)})]{} N^{(1)}$$

xPG

Suppose the number of proposed paths is $M = 4$. On the time interval $[(k-1)\tau, k\tau]$, given $X^{(0:4)}_{(k-1)\tau}$ and θ, we compute the Poisson means of all paths and order them as:

$$\mu^{(2)} < \mu^{(3)} < \mu^{(0)} < \mu^{(4)} < \mu^{(1)}$$

where $\mu^{(0)}$ is the Poisson mean of the reference path.

For each reaction type, given the number of such events in the reference path, $N^{(0)}$, simulate the corresponding number of events in the proposed paths.

$$N^{(2)} \leftarrow N^{(3)} \leftarrow N^{(0)} \xrightarrow[+\text{Pois}(\mu^{(4)}-\mu^{(0)})]{} N^{(4)} \xrightarrow[+\text{Pois}(\mu^{(1)}-\mu^{(4)})]{} N^{(1)}$$

$$N^{(2)} \xleftarrow[\text{Bin}\left(N^{(3)}, \frac{\mu^{(2)}}{\mu^{(3)}}\right)]{} N^{(3)} \xleftarrow[\text{Bin}\left(N^{(0)}, \frac{\mu^{(3)}}{\mu^{(0)}}\right)]{} N^{(0)} \rightarrow N^{(4)} \rightarrow N^{(1)}$$

xPG

$N(t)$ is a Poisson process with rate 1.

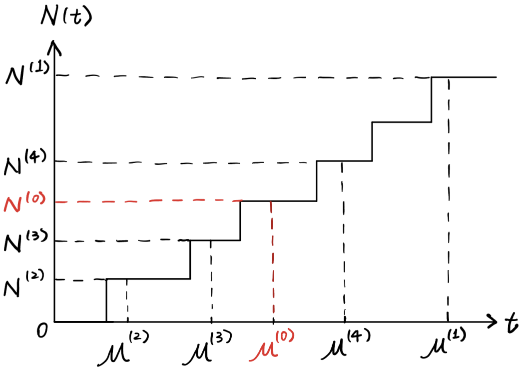


Figure 17: Sample path of a Poisson process $N(t)$ with rate 1

xPG

We introduce a tuning parameter δ to control the correlation between the proposed paths and the reference path. On the interval $[(k-1)\tau, k\tau]$, suppose the Poisson means satisfy

$$\mu^{(2)} < \mu^{(3)} < \mu^{(0)} < \mu^{(4)} < \mu^{(1)},$$

and let $N^{(0)}$ be the number of reaction events in the reference path. Then we proceed as follows:

xPG

We introduce a tuning parameter δ to control the correlation between the proposed paths and the reference path. On the interval $[(k-1)\tau, k\tau]$, suppose the Poisson means satisfy

$$\mu^{(2)} < \mu^{(3)} < \mu^{(0)} < \mu^{(4)} < \mu^{(1)},$$

and let $N^{(0)}$ be the number of reaction events in the reference path. Then we proceed as follows:

1. $\tilde{N}^{(0)} \sim \text{Bin}\left(N^{(0)}, 1-\delta\right)$

xPG

We introduce a tuning parameter δ to control the correlation between the proposed paths and the reference path. On the interval $[(k-1)\tau, k\tau]$, suppose the Poisson means satisfy

$$\mu^{(2)} < \mu^{(3)} < \mu^{(0)} < \mu^{(4)} < \mu^{(1)},$$

and let $N^{(0)}$ be the number of reaction events in the reference path. Then we proceed as follows:

1. $\tilde{N}^{(0)} \sim \text{Bin}\left(N^{(0)}, 1-\delta\right)$

2. Simulate $\tilde{N}^{(m)}$, $m = 1, \ldots 4$ using Binomial thing and Poisson update sequentially

$$\tilde{N}^{(2)} \leftarrow \tilde{N}^{(3)} \leftarrow \tilde{N}^{(0)} \xrightarrow[+\text{Pois}\left((1-\delta)(\mu^{(4)}-\mu^{(0)})\right)]{} \tilde{N}^{(4)} \xrightarrow[+\text{Pois}\left((1-\delta)(\mu^{(1)}-\mu^{(4)})\right)]{} \tilde{N}^{(1)}$$

$$\tilde{N}^{(2)} \xleftarrow[\text{Bin}\left(\tilde{N}^{(3)}, \frac{\mu^{(2)}}{\mu^{(3)}}\right)]{} \tilde{N}^{(3)} \xleftarrow[\text{Bin}\left(\tilde{N}^{(0)}, \frac{\mu^{(3)}}{\mu^{(0)}}\right)]{} \tilde{N}^{(0)} \rightarrow \tilde{N}^{(4)} \rightarrow \tilde{N}^{(1)}$$

xPG

We introduce a tuning parameter δ to control the correlation between the proposed paths and the reference path. On the interval $[(k-1)\tau, k\tau]$, suppose the Poisson means satisfy

$$\mu^{(2)} < \mu^{(3)} < \mu^{(0)} < \mu^{(4)} < \mu^{(1)},$$

and let $N^{(0)}$ be the number of reaction events in the reference path. Then we proceed as follows:

1. $\tilde{N}^{(0)} \sim \text{Bin}\left(N^{(0)}, 1-\delta\right)$

2. Simulate $\tilde{N}^{(m)}$, $m = 1, \ldots 4$ using Binomial thing and Poisson update sequentially

$$\tilde{N}^{(2)} \leftarrow \tilde{N}^{(3)} \leftarrow \tilde{N}^{(0)} \xrightarrow[+\text{Pois}\left((1-\delta)(\mu^{(4)}-\mu^{(0)})\right)]{} \tilde{N}^{(4)} \xrightarrow[+\text{Pois}\left((1-\delta)(\mu^{(1)}-\mu^{(4)})\right)]{} \tilde{N}^{(1)}$$

$$\tilde{N}^{(2)} \xleftarrow[\text{Bin}\left(\tilde{N}^{(3)}, \frac{\mu^{(2)}}{\mu^{(3)}}\right)]{} \tilde{N}^{(3)} \xleftarrow[\text{Bin}\left(\tilde{N}^{(0)}, \frac{\mu^{(3)}}{\mu^{(0)}}\right)]{} \tilde{N}^{(0)} \rightarrow \tilde{N}^{(4)} \rightarrow \tilde{N}^{(1)}$$

3. $N^{(m)} = \tilde{N}^{(m)} + \text{Pois}\left(\delta\mu^{(m)}\right)$, $m = 1, \ldots, 4$.

Experiment results

SIS model

$$\text{Susceptible} \underset{\gamma I}{\overset{\beta SI}{\rightleftharpoons}} \text{Infectious}$$

$$\text{Reaction 1 } (R_1) : S + I \xrightarrow{\beta SI} 2I,$$

$$\text{Reaction 2 } (R_2) : I \xrightarrow{\gamma I} S \quad ,$$

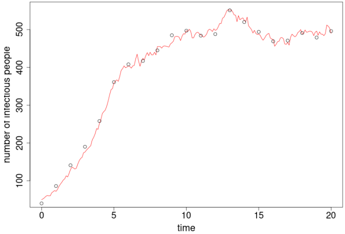


Figure 18: One simulated trajectory of the number of infectious individuals in the SIS model with 21 observations, using $\beta = 1.2$ and $\gamma = 0.6$

SIS models

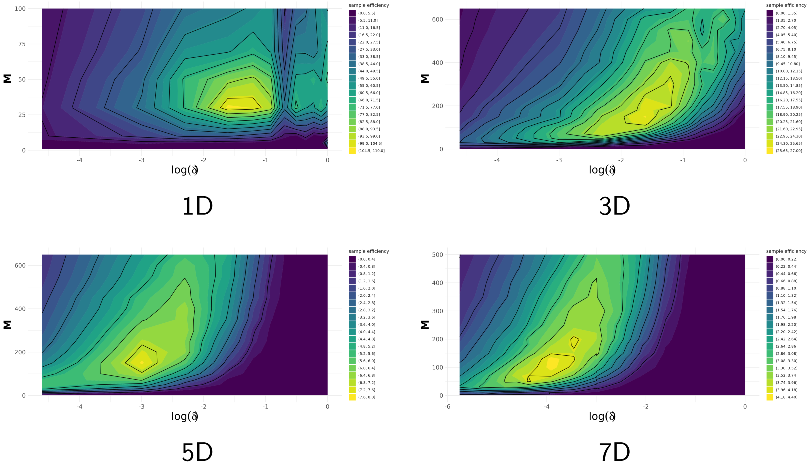


Figure 19: Contour plots of $\mathrm{ESS}(X_0)/M$ after 10^5 iterations in 1D, 3D, 5D, and 7D; latent states are products of SIS models with shared parameters $\beta = 1.2$ and $\gamma = 0.6$.

SIR models

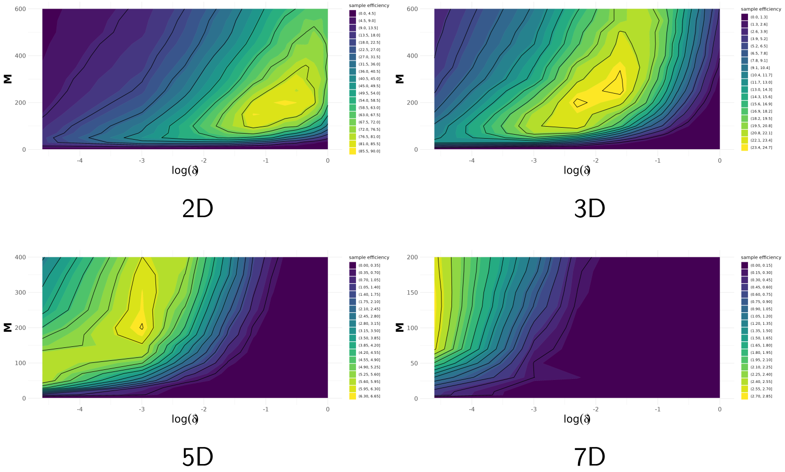


Figure 20: Contour plots of the $\text{ESS}(X_0)/M$ after 10^5 iterations in 2D, 3D, 5D and 7D; the true latent states are product of SIR models with shared model parameters $\beta = 0.6$ and $\gamma = 0.2$.

Autoregulatory model

The total number of copies of DNA, G, is fixed, and the reactions are:

$$\text{DNA} + P_2 \xrightarrow{\theta_1 (G - X_4) X_3} \text{DNA} \cdot P_2,$$

$$\text{DNA} \xrightarrow{\theta_3 (G - X_4)} \text{DNA} + \text{RNA},$$

$$\text{DNA} \cdot P_2 \xrightarrow{\theta_2 X_4} \text{DNA} + P_2,$$

$$\text{RNA} \xrightarrow{\theta_4 X_1} \text{RNA} + P,$$

$$2P \xrightarrow{\theta_5 X_2 (X_2 - 1)/2} P_2,$$

$$\text{RNA} \xrightarrow{\theta_7 X_1} \varnothing,$$

$$P_2 \xrightarrow{\theta_6 X_3} 2P, \quad P \xrightarrow{\theta_8 X_2} \varnothing,$$

where X_1, X_2, X_3, X_4 denote the counts of RNA, P, P_2, and DNA $\cdot P_2$, respectively.

Autoregulatory model

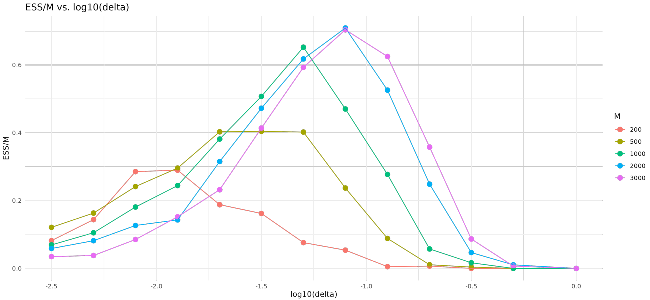


Figure 21: The efficiency of xPG again $\log_{10}\delta$ after 2×10^5 iterations, with each curve corresponding to a different value of M, $M = 200, 500, 1000, 2000, 3000$

Tuning parameters M and δ

Tuning Parameters M and δ

- α_{ref}: expected probability of accepting a path that has not coalesced with the reference path at time 0.
- α_{val}: expected probability of accepting a new value for X_0.

Key Idea

$$\text{ESS}(X_0) \approx (\text{expected squared jump distance moved}) \times \alpha_{\text{val}}.$$

For xPG, when δ is fixed,

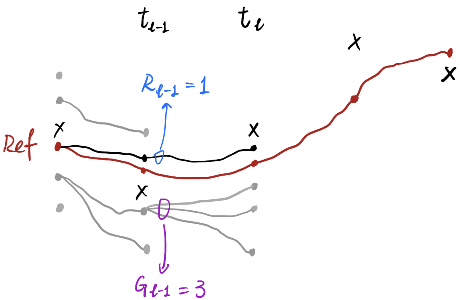
$$\text{Eff} = \frac{\text{ESS}(X_0)}{M} \propto \frac{\alpha_{\text{val}}}{M}.$$

Tuning parameters M and δ

- **Good particle:** its ancestor at time 0 has a value different from $x_0^{(0)}$
- **Bad particle:** its ancestor at time 0 has value $x_0^{(0)}$.

Let G_l be the number of good particles **after resampling** at time t_l, and let R_l be the number of bad particles. Since all particles are either good or bad,

$$G_l + R_l = M.$$



Tuning parameters M and δ

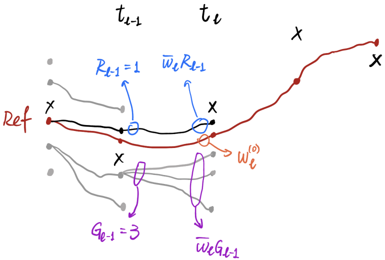
Assumption 1. For $l = 1, \dots, L$, we assume that

$$\frac{\sum_{i \in \{\text{time 0 ancestor}=0\}} w_l^i}{R_{l-1}} = \bar{w}_l = \frac{\sum_{i \in \{\text{time 0 ancestor}\neq 0\}} w_l^i}{G_{l-1}}, \qquad (3)$$

where $\bar{w}_l := \frac{1}{M} \sum_{m=1}^{M} w_\ell^{(m)}$ denotes the average weight of the proposed particles computed before the resampling at time t_l.

Assumption 2. G_{l-1} and R_{l-1} are independent of $W_l^{(0)}$.

Based on these two assumptions, the tuning suggestion is to tune the acceptance rate of the initial state α_{ref} to a value of 0.368.



Tuning parameters M and δ

Target: $\alpha_{\text{val}} \approx \frac{\mathbb{E}(G_L)}{M+1}$

Let G_{-1} be the number of good particles that are initially proposed particles at time zero. Its expectation is given by

$$\mathbb{E}(G_{-1}) = M(1 - p_*^\delta),$$

where p_*^δ denotes the probability of proposing, at time zero, a particle that has the same value as the initial state of the reference path. This probability depends on δ; larger values of δ correspond to smaller p_*^δ.

Tuning parameters M and δ

Therefore, at the observation times, the expected number of particles whose ancestor at time zero has a different value from $X_0^{(0)}$ is given by

$$\mathbb{E}(G_l) = M \times \mathbb{E}\left(\frac{\bar{w}_l G_{l-1}}{\bar{w}_l M + w_l^{(0)}}\right) \approx \mathbb{E}(G_{l-1}) \times \frac{\mathbb{E}(\bar{w}_l)M}{\mathbb{E}(\bar{w}_l)M + \mathbb{E}(w_l^0)}, \quad l = 0, 1, \ldots, L \tag{4}$$

where $w_l^{(0)}$ is the weight of the particle in the reference path at the observation time t_l. The approximations are obtained from the *strong law of large numbers*.

Tuning parameters M and δ

$$
\begin{aligned}
\mathbb{E}(G_L) &\approx (1-p_*^\delta)M \times \prod_{l=0}^{L} \frac{\mathbb{E}(\bar{w}_l)M}{\mathbb{E}(\bar{w}_l)M + \mathbb{E}(w_l^0)} \\
&= (1-p_*^\delta)M \times \prod_{l=0}^{L} \left(1 - \frac{\mathbb{E}(w_l^0)}{\mathbb{E}(\bar{w}_l)M + \mathbb{E}(w_l^0)}\right) \\
&\approx (1-p_*^\delta)M \times \exp\left(-\frac{1}{M}\sum_{l=0}^{L} \frac{\mathbb{E}(w_l^0)}{\mathbb{E}(\bar{w}_l)}\right) \\
&= (1-p_*^\delta)M \times \exp\left(-\frac{H}{M}\right),
\end{aligned} \tag{5}
$$

where $H = \sum_{l=0}^{L} \frac{\mathbb{E}(w_l^0)}{\mathbb{E}(\bar{w}_l)}$.

Tuning parameters M and δ

Therefore, the acceptance rate α_{val} is approximated by

$$\alpha_{\text{val}} = \frac{\mathbb{E}(G_L)}{M+1} = (1 - p_*^\delta)\frac{M}{M+1}\exp\left(-\frac{H}{M}\right). \tag{6}$$

and the actual acceptance rate α_{ref} is given by

$$\alpha_{\text{ref}} = \frac{M}{M+1}\exp(-\frac{H}{M}), \tag{7}$$

If we fix δ for xPG, the efficiency of xPG is proportional to

$$\text{Eff} \propto \frac{\alpha_{\text{val}}}{M} \approx (1 - p_*^\delta)\frac{1}{M}\exp(-\frac{H}{M}).$$

Take derivative of Eff with respect to M, we obtain $\hat{M} = H$. The optimal actual acceptance rate α_{ref} is given by

$$\hat{\alpha}_{\text{ref}} = e^{-1} \approx 0.368.$$

Tuning parameters M and δ

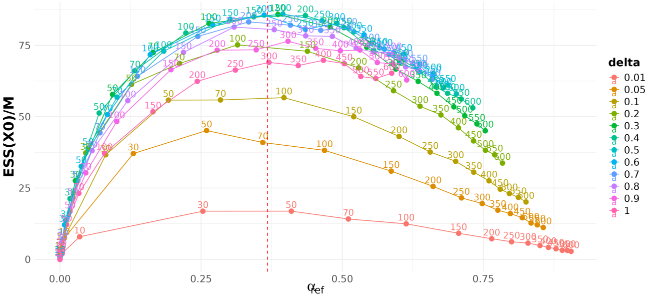


Figure 24: $\mathrm{ESS}(X_0)/M$ against α_{ref}; the vertical red dashed line represents $\alpha_{\text{ref}} = 0.368$; the true latent process is the product of two independent SIR models with shared model parameters $\beta = 0.6$ and $\gamma = 0.2$.

Tuning strategy

For any $\delta \in [0, 1]$, we choose a large $M = M^*$ such as $M^* = 100$ or $M^* = 1000$ and run the algorithm for a moderate number of iterations, noting the actual acceptance rate $\alpha_{\mathrm{ref}}(\delta, M^*)$. We may then derive the optimal M for δ, denoted by $\hat{M}_\delta$, from equation 7 by

$$\hat{M}_\delta = -M^* \log\left(\frac{M^* + 1}{M^*}\alpha_{\mathrm{ref}}(\delta, M^*)\right). \tag{8}$$

Tuning strategy

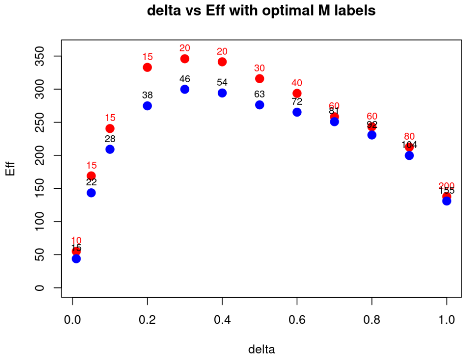


Figure 25: The blue dots are $\delta \in \{0.01, 0.05, (1:10) \times 0.1\}$ versus $\text{Eff} = \text{ESS}(X_0)/\hat{M}_\delta$ with the $\hat{M}_\delta$ labeled; $\text{ESS}(X_0)$ is obtained by running xPG with $(\delta, \hat{M}_\delta)$; the red dots represent the maximum efficiency attained for each δ, with the corresponding value of M (that achieves this maximum) shown as the label

Additional Work and Future Directions

Additional work completed:

- Derived an xPG algorithm for reaction networks based on **exact simulation** of the MJPs
- Found a way to apply **ancestor sampling** for the tau-leap model in some cases to enhance particle diversity.
- Apply the proposed methods to multi-dimensional state space systems where correlations exist between states across dimensions

Future research directions:

- Extend the methodology to discrete-time **chain-binomial** epidemic model

References

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**(3): 269–342.
URL: *http://dx.doi.org/10.1111/j.1467-9868.2009.00736.x*

Corenflos, A. and Finke, A. (2024). Particle-mala and particle-mgrad: Gradient-based mcmc methods for high-dimensional state-space models.
URL: *https://arxiv.org/abs/2401.14868*

Finke, A. and Thiery, A. H. (2023). Conditional sequential monte carlo in high dimensions, The Annals of Statistics **51**(2).
URL: *http://dx.doi.org/10.1214/22-AOS2252*

Lindsten, F., Jordan, M. I. and Schön, T. B. (2014). Particle gibbs with ancestor sampling.
URL: *https://arxiv.org/abs/1401.0604*

Malory, S. J. (2021). Bayesian inference for stochastic processes, Lancaster University (United Kingdom).

Particle Filter

Given:

- A sequence of L observations: $(y_{t_1}, \ldots, y_{t_L})$
- Known model parameters θ

Goal:

- Infer the latent state X_{t_i}, $i = 1, \ldots, L$, given $y_{t_1}, \ldots, y_{t_i}$.
- Target distribution:

$$p(X_{t_i} \mid y_{t_{1:i}}, \theta) \quad i = 1, \ldots, L$$

Particle Filter (PF)

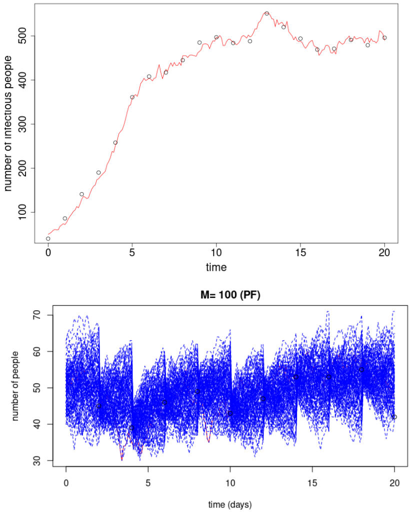


Figure 26: A single run of PF on the SIS model, with $M = 100$, $\lambda = 0.8$ and $\mu = 0.4$.

Conditional PF within the Framework of the τ-Leap

Algorithm 2 conditional PF within the Framework of the τ-Leap

Require: Observations $y = \left(y_{K_1\tau}, y_{K_2\tau}, \ldots, y_{K_L\tau}\right)$ and reference state process $\left(x_0^{(0)}, n_\tau^{(0)}, n_{2\tau}^{(0)}, \ldots, n_{K_L\tau}^{(0)}\right)$

1: **Initialize:** Simulate $x_0^{(1)}, \ldots, x_0^{(M)}$ based on $x_0^{(0)}$ such that $P_0(x_0^{(0)})P(x_0^{(1:M)}|x_0^{(0)}) = P_0(x_0^{(i)})P(x_0^{(-i)}|x_0^{(i)})$, $i = 1, \ldots, M$

2: **for** $k = 1, \ldots, K_L$ **do**

3: Given $x_{(k-1)\tau}^{(0)}, x_{(k-1)\tau}^{(1)}, \ldots, x_{(k-1)\tau}^{(M)}$ and $n_{k\tau}^{(0)}$, simulate $n_{k\tau}^{(1)}, \ldots, n_{k\tau}^{(M)}$

4: Set $j = 1$

5: **if** $k = K_j$ **then**

6: Resample M particles from $\left\{x_{k\tau}^{(0)}, x_{k\tau}^{(1)}, \ldots, x_{k\tau}^{(M)}\right\}$. The weight of particle $x_{k\tau}^{(i)}$ is proportional to the likelihood $g(y_{K_j\tau}|x_{k\tau}^{(i)})$, $i = 0, 1, \ldots, M$

7: Replace $\left\{x_{k\tau}^{(1)}, \ldots, x_{k\tau}^{(M)}\right\}$ with the resampled particles

8: Set $j = j + 1$

9: **end if**

10: **end for**

11: **return** $(M + 1)$ state processes

Validity of One-step xPGibbs

Imagine we now have an observation at y_1 with a likelihood of $f(y_1|x_{K\tau})$. We have a reference path, which is $x_0^{(0)}$ and $x_{(1:K)\tau}^{(0)}$. From these, we can simulate exchangeable $X_0^{1:M}$ and $N_{1:K}^{1:M}$. We accept $x_{K\tau}^{(i)}$ with a probability of

$$\alpha(0,i) = \frac{f(y_1|x_{K\tau}^{(i)})}{\sum_{j=1}^{M} f(y_1|x_{K\tau}^{(j)})}.$$

If $N_{1:K}^{(0)}$ arises from their joint posterior then they have a mass function proportional to

$$f(y_1|x_{K\tau}^{(0)})\mathbb{P}\left(X_0^{(0)} = x_0^{(0)}\right)\prod_{k=1}^{K}\mathbb{P}\left(N_k^{(0)} = n_k^{(0)}|x_{(k-1)\tau}^{(0)}\right),$$

where, for $k \geq 2$, $x_{(k-1)\tau}^{(0)}$ is a function of $x_{(k-2)\tau}^{(0)}$ and $n_{k-1}^{(0)}$.

Validity of One-step xPGibbs

The probability of proposing all of the other random variables is

$$\mathbb{P}(X_0^{(1:M)} = x_0^{(1:M)} | x_0^{(0)}) \prod_{k=1}^{K} \mathbb{P}\left(\tilde{N}_k^{(0)} = \tilde{n}_k^{(0)}, \tilde{N}_k^{(1:M)} = \tilde{n}_k^{(1:M)}, N_k^{(1:M)} = n_k^{(1:M)} | n_k^{(0)}, x_{(k-1)\tau}^{(0:M)}\right).$$

The product of the posterior mass function and the proposal mass function can be re-written as

$$f(y_1 | x_{K\tau}^{(0)}) \mathbb{P}\left(X_0^{(i)} = x_0^{(i)}\right) \mathbb{P}\left(X_0^{(-i)} = x_0^{(-i)} | X_0^{(i)} = x_0^{(i)}\right)$$
$$\times \prod_{k=1}^{K} \mathbb{P}\left(N_k^{(i)} = n_k^{(i)}, \tilde{N}_k^{(i)} = \tilde{n}_k^{(i)}, \tilde{N}_k^{(-i)} = \tilde{n}_k^{(-i)}, N_k^{(-i)} = n_k^{(-i)} | x_{(k-1)\tau}^{(0:M)}\right)$$

Multiplying this by $\alpha(0, i)$, where $i \in \{0, \ldots, M\}$, gives the probability of starting from the i-th path, proposing M other paths, and then accepting path 0.