Analysis of Markovian Population Models

Dissertation Defense

Michael Backenköhler

Saarland Informatics Campus

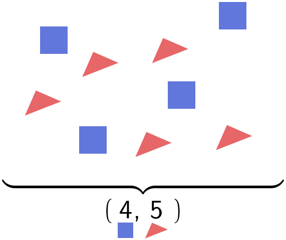
September 20, 2022

Motivation

- stochastic concentration changes
- systems biology: switches,
- applications: queueing, metabolic networks, bio-switches, traffic etc.
- goal: do *rigorous* analysis on such models

Markovian Population Models

Semantics



- state space ∼ population sizes
- often huge to infinite

- continuous time
- exponential jump times / CTMC dynamics

Markovian Population Models

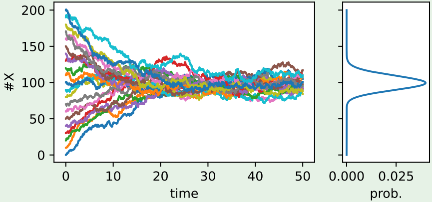
Stationary Distribution – Foster-Lyapunov Functions

- ergodic chains converge to unique distribution ($t \to \infty$)

Example

$$X \xrightarrow{0.1\cdot\#X} \varnothing, \quad \alpha_1(x) = 0.1x$$

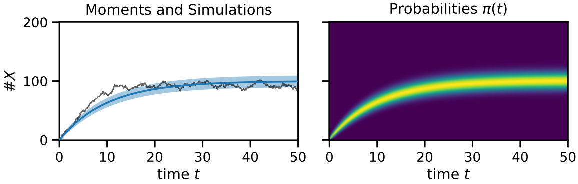
$$\varnothing \xrightarrow{10} X, \quad \alpha_2(x) = 10$$



- how does this distribution look like for infinite state-spaces?
- use Foster-Lyapunov function to bound sets
- locally augment functions for tighter sets / bounds

Markovian Population Models

Moment Dynamics



Moment formula

Moments such as mean $\mathrm{E}\left(X_t\right)$ and variance $\mathrm{E}\left(X_t^2\right) - \mathrm{E}\left(X_t\right)^2$ are described by (often linear) ODEs.

$$\frac{d}{dt}\mathrm{E}\left(f(\vec{X}_t)\right) = \sum_{j=1}^{n_R} \mathrm{E}\left(\left(f(\vec{X}_t + \vec{\nu}_j) - f(\vec{X}_t)\right)\alpha_j(\vec{X}_t)\right)$$

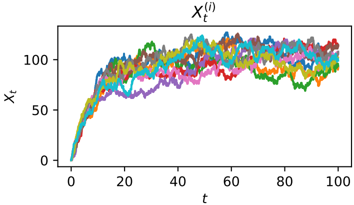
Markovian Population Models

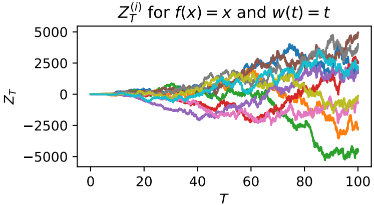
Martingale Process

- multiply time-weighting: $w(t) = t^k$, $k \in \mathbb{N}$ or $w(t) = \exp(\lambda t)$
- analytic integration and resulting martingale process

$$Z_T \coloneqq w(T) f(\vec{X}_T) - w(0) f(\vec{X}_0) - \int_0^T \frac{dw(t)}{dt} f(\vec{X}_t)\, dt$$
$$- \sum_{j=1}^{n_R} \int_0^T w(t) \left(f(\vec{X}_t + \vec{v}_j) - f(\vec{X}_t) \right) \alpha_j(\vec{X}_t)\, dt$$

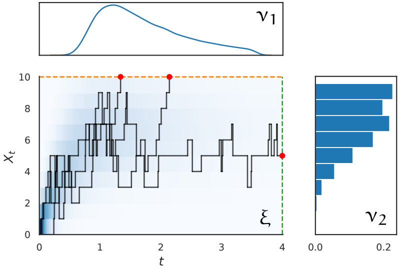
- known expectation: $E(Z_T) = 0$, $\forall T \geqslant 0$





Bounding Mean First-Passage Times

Linear Moment Constraints



- $\tau = \inf\{X_t \geqslant H \mid t \geqslant 0\} \wedge T$
- exp. occupation measure ξ
- exit location measures ν_1, ν_2

$$0 = \mathrm{E}\left(Z_T\right) = T^k \overbrace{\mathrm{E}\left(X_\tau^m; \tau = T\right)}^{\nu_1} + H^m \overbrace{\mathrm{E}\left(\tau^k; \tau < T, X_\tau = H\right)}^{\nu_2}$$
$$- 0^k x_0^m + \sum_i c_i \underbrace{\mathrm{E}\left(\int_0^\tau t^{k_i} X_t^{m_i}\, dt\right)}_{\xi}$$

[BBW20]

Moment Matrices

The *moment matrix* mus be *positive semi-definite*.

$$\mathrm{E}\begin{pmatrix} X^0 & X^1 & X^2 & \dots & X^n \\ X^1 & X^2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ X^n & X^{n+1} & \dots & \dots & X^{2n} \end{pmatrix} \succeq 0\,,$$

where $M \succeq 0$ iff. $\forall v \in \mathbb{R}^n . v^\mathsf{T} M v \geqslant 0$.

Example

Let $M = \begin{pmatrix} 1 & \mathrm{E}(X) \\ \mathrm{E}(X) & \mathrm{E}\left(X^2\right) \end{pmatrix}$. Then $\det M = \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2 = \sigma^2 \geqslant 0$.

[BBW20]

- measure support can be restricted using semi-definite constraints
- resulting SDPs can be solved using off-the-shelf software.

Semi-Definite Program (SDP)

An optimization problem with

1. *linear constraints* on moments and
2. *positive semi-definite constraints* on certain matrices.

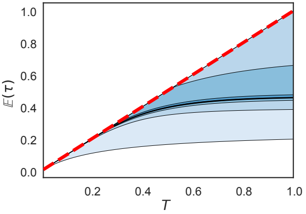
- alternative: linear *Hausdorff constraints* instead of semi-definite constraints

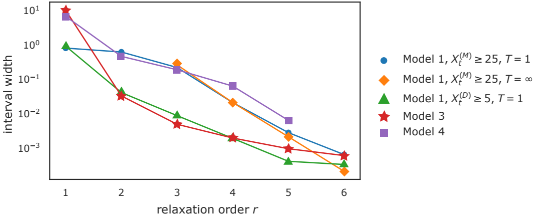
$$\int_{[0,1]^{\vec{n}}} \vec{x}^{\vec{\ell}} (1 - \vec{x})^{\vec{k}} \, d\mu(\vec{x}) \geqslant 0$$

[BBW20]

Bounding Mean First-Passage Times

Results



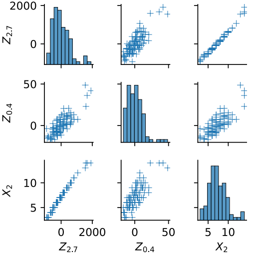


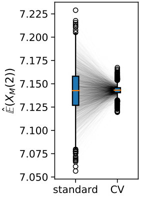
- fast convergence of bounds with increasing order
- SDPs are usually solved within seconds
- numerically challenging (inherent stiffness)
- scaling state-space / model size is difficult

[BBW20]

Linear Control Variates

Using Correlated RVs with Known Expected Value





- improve MC estimates using Z_T
- use correlations between Z_T and X_T
- $E(X_T + bZ_T)$ instead of $E(X_T)$ (recall $E(Z_T) = 0$)

Linear Control Variates

Given a control variate vector $\vec{Z}$, the estimator

$$\hat{V} - (\hat{\Sigma}_{\vec{Z}}^{-1}\hat{\Sigma}_{\vec{Z}V})^{\mathsf{T}}\hat{\vec{Z}}$$

has lower or equal variance as $\hat{V}$.

[BBW19]

Linear Control Variates

Efficiency Trade-off

cost: slowdown

$$c_{\text{old}} \big/ c_{\text{new}}$$

- computing $\int_0^T w(t) X_t^m \, dt$
- computing the estimate

benefit: variance reduction

$$\sigma^2_{\text{new}} \big/ \sigma^2_{\text{old}}$$

- highly correlated variates

Approach: Assess correlations between k candidates and the target RV V

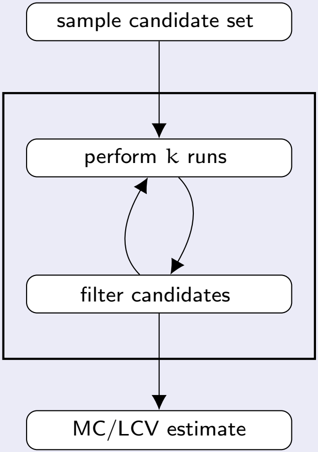
$$\begin{bmatrix} 1 & \dots & \rho_{1k} & \rho_{1v} \\ \vdots & \ddots & \vdots & \vdots \\ \rho_{k1} & \dots & 1 & \rho_{kv} \\ \rho_{v1} & \dots & \rho_{vk} & 1 \end{bmatrix}.$$

[BBW19; BBW22]

Linear Control Variates

Selection by Filtering

Algorithm 1



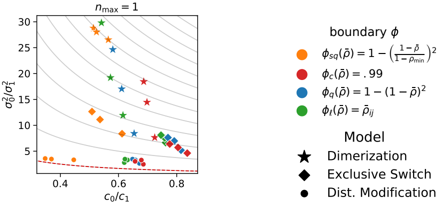
[BBW19]

Filter criteria:

1. low target correlation

$$\rho_{i\nu} < \max\left(0.1, \frac{\max_j \rho_{j\nu}}{k_{\min}}\right)$$

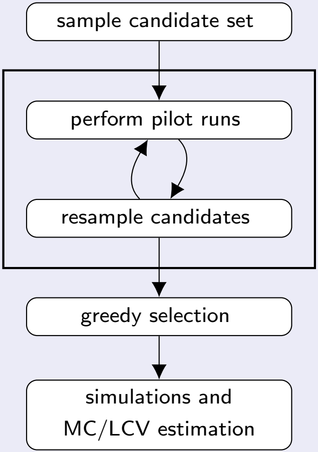
2. various redundancy heuristics



Linear Control Variates

Selection by Resampling

Algorithm 2

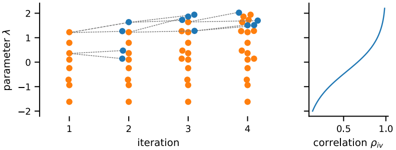


[BBW22]

Filter criteria:

1. *resampling* proportional to improvement

$$\gamma_{k\nu} = (1 - \rho_{k\nu}^2)^{-1}$$



2. *selection* by best improvement of one CV

$$\underset{1 \leqslant i \leqslant |P_{\text{all}}|}{\arg\max}\; \hat{\gamma}_{i\nu} \prod_{\substack{1 \leqslant j \leqslant |P_{\text{all}}| \\ (m_j, \lambda_j) \in P^*}} \hat{\gamma}_{ij}^{-1}$$

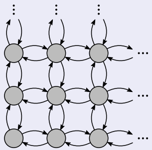
Linear Control Variates

Results

- best example?

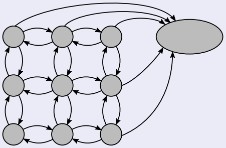
Finite-Space Projection

Original



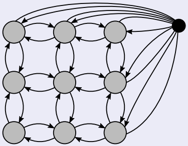
- infinite

Sink state



- transient analysis
- keep track of approx. error

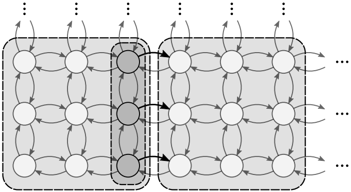
Redirection



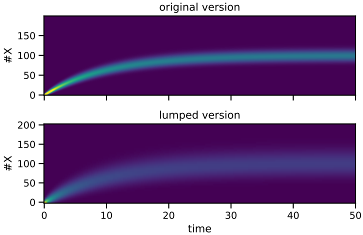
- stationary dist.
- dependent on redirection

State-Space Aggregation

Treating Hyper-Cubes of States as One



- hyper-cube macro-states
- *assumption*: uniform dist. within
- closed-form transition rates



- resulting distribution more “flat”
- locate main probability mass

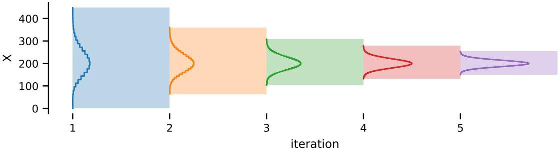
[Bac+21a]

Stationary Distribution

Iterative Refinement Algorithm

A simple refinement based on approximate solutions:

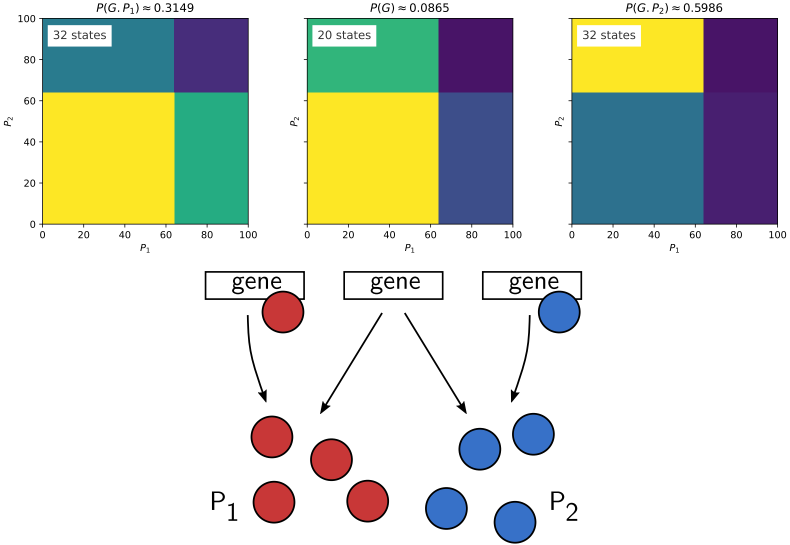
1. start with macro-states of size 2^k
2. compute approximate distribution
3. remove states with low probability
4. split the remaining states
5. go to step 2



[Bac+21b; Bac+21a]

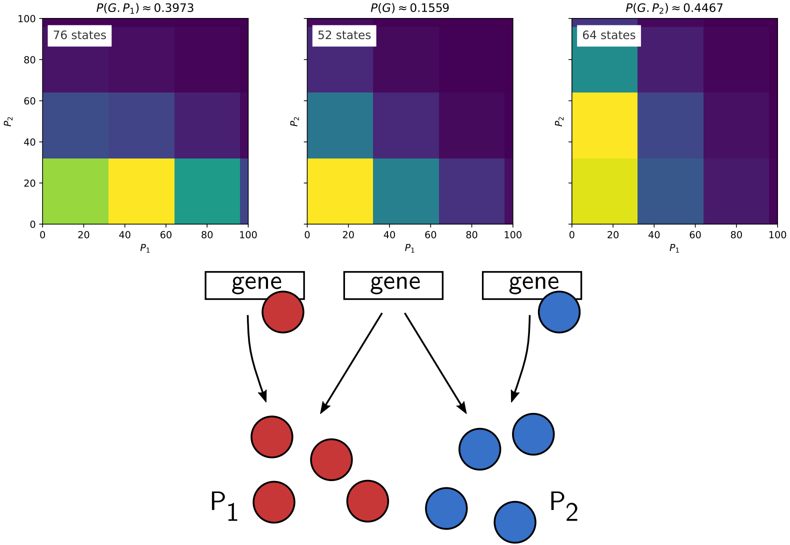
Stationary Distribution

Examples



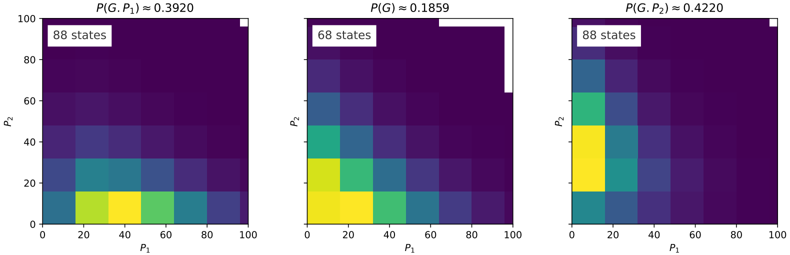
Stationary Distribution

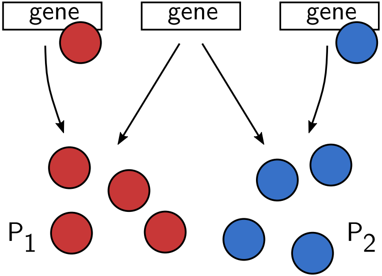
Examples



Stationary Distribution

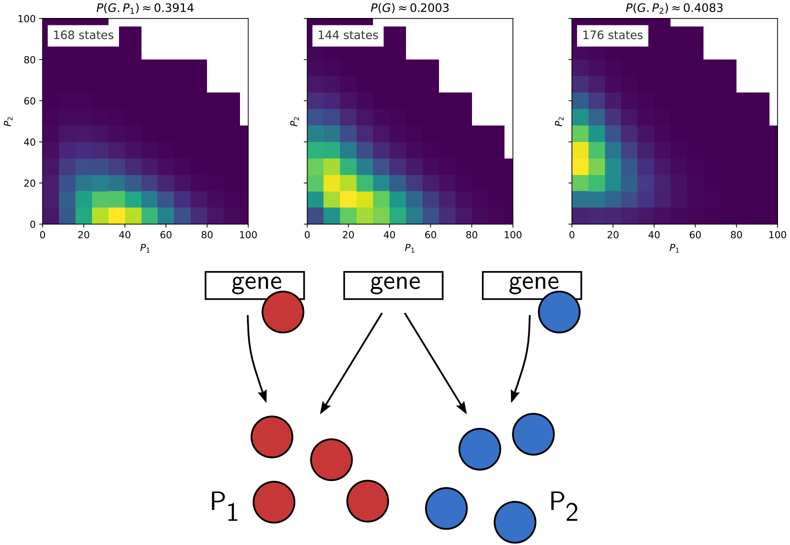
Examples





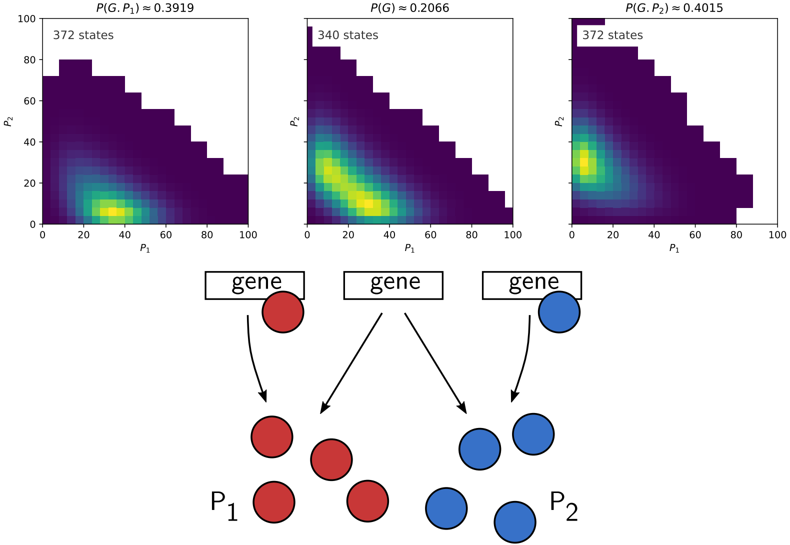
Stationary Distribution

Examples



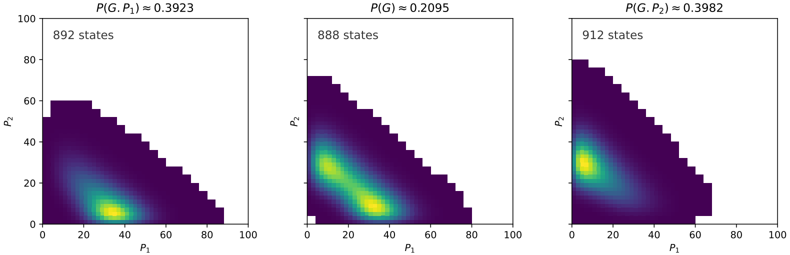
Stationary Distribution

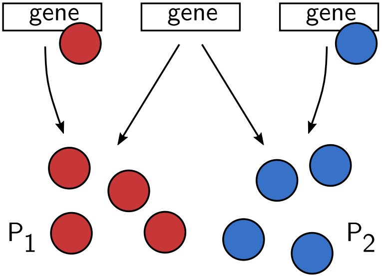
Examples



Stationary Distribution

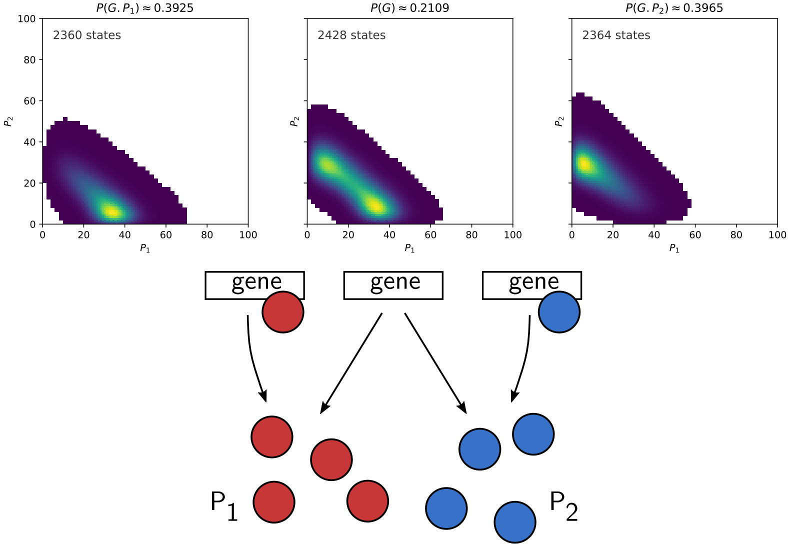
Examples





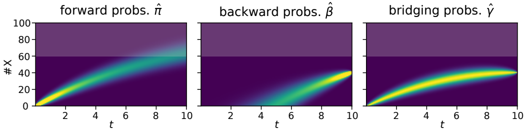
Stationary Distribution

Examples



Bridging Problem

Dynamical Analysis Under Initial *and* Terminal Constraints



Forward Probabilities π

How the process evolves with time: $\Pr(X_t = x \mid X_0 = 0)$

Backward Probabilities β

Probability of ending up in a given state: $\Pr(X_T = 40 \mid X_t = x)$

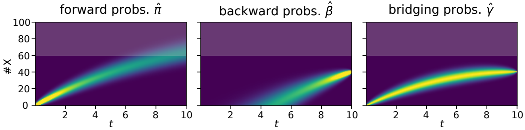
Bridging Probabilities γ

In between: $\Pr(X_t = x \mid X_0 = 0, X_T = 40)$

[Bac+21b]

Bridging Problem

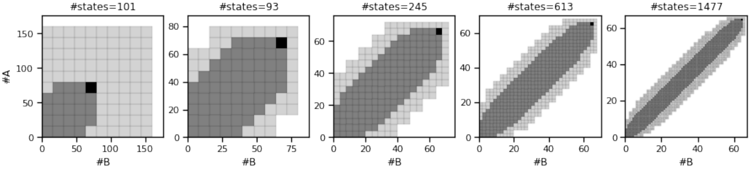
Refinement



- bridging distribution:

$$\gamma(x_i, t) = \pi(x_i, t)\beta(x_i, t)/\pi(x_g, T)$$

- record intermediary times
- remove or split based on $\hat{\gamma}(x_i, t)$



[Bac+21b]

Bridging Problem

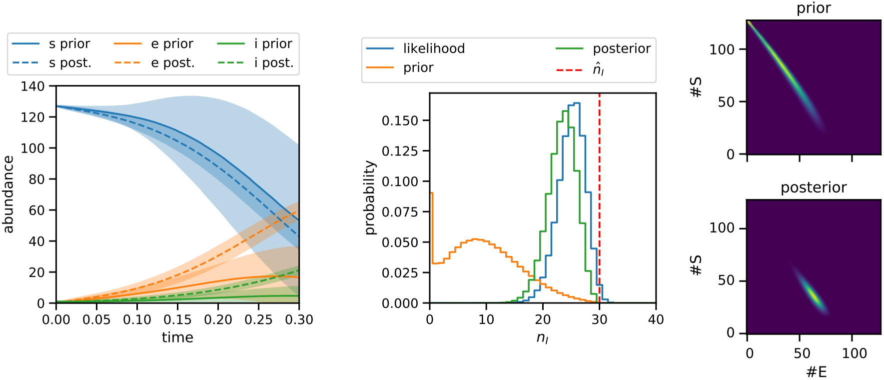
Bayesian Filtering in an SEIR model

$$\mathrm{S} + \mathrm{I} \to \mathrm{E} + \mathrm{I} \qquad \mathrm{E} \to \mathrm{I} \qquad \mathrm{I} \to \mathrm{R}$$

We know:

- initial state
- uncertain measurement of I at $\mathrm{T} = 0.3$

We are interested in the posterior at T.



[Bac+21b]

Contributions

- local augmentation of Foster-Lyapunov functions
- bounding of mean first-passage times [BBW20]
- variance reduction for MC estimation [BBW19; BBW22]
- state-space aggregation scheme
  - stationary distribution [Bac+21a]
  - bridging distribution [Bac+21b]
  - importance sampling

References I

[BBW20] Michael Backenköhler, Luca Bortolussi, and Verena Wolf. "Bounding Mean First Passage Times in Population Continuous-Time Markov Chains". In: *17th International Conference on Quantitative Evaluation of SysTems*. Vol. 12289. Lecture Notes in Computer Science. Springer, 2020, pp. 155–174.

[BBW19] Michael Backenköhler, Luca Bortolussi, and Verena Wolf. "Control Variates for Stochastic Simulation of Chemical Reaction Networks". In: *17th International Conference on Computational Methods in Systems Biology*. Vol. 11773. Lecture Notes in Computer Science. Springer, 2019, pp. 42–59.

[BBW22] Michael Backenköhler, Luca Bortolussi, and Verena Wolf. "Variance Reduction in Stochastic Reaction Networks using Control Variates". In: *Principles of Systems Design – Essays Dedicated to Thomas A. Henzinger on the Occasion of His 60th Birthday*. Vol. 13660. Lecture Notes in Computer Science. Springer, 2022.

References II

[Bac+21a] Michael Backenköhler et al. "Abstraction-Guided Truncations for Stationary Distributions of Markov Population Models". In: *18th International Conference on Quantitative Evaluation of SysTems*. Vol. 12846. Lecture Notes in Computer Science. Springer, 2021, pp. 351–371.

[Bac+21b] Michael Backenköhler et al. "Analysis of Markov Jump Processes under Terminal Constraints". In: *27th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Vol. 12651. Lecture Notes in Computer Science. Springer, 2021, pp. 210–229.