

# Analysis of Markovian Population Models

## Dissertation Defense

Michael Backenköhler

Saarland Informatics Campus

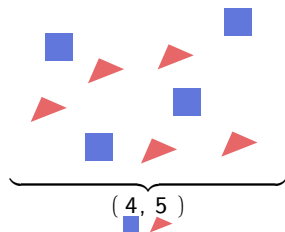
September 20, 2022

# Motivation

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

# Markovian population models

## Framework



- ▶ populations of identical agents
  - ▶ state space  $\sim$  population sizes
  - ▶ often huge to infinite
- 
- ▶ continuous time
  - ▶ exponential jump times / CTMC dynamics
  - ▶ Kolmogorov equation for probabilities:

$$\frac{d}{dt}\pi(t) = \pi(t)Q$$

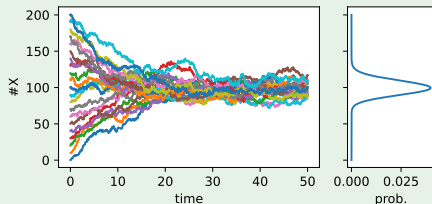
# Markovian population models

How they behave

## Example (birth-death process)

$$X \rightarrow \emptyset, \quad \alpha_1(x) = 0.1 \cdot x$$

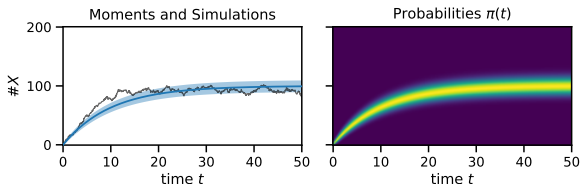
$$\emptyset \rightarrow X, \quad \alpha_2(x) = 10$$



- ▶ changes depend on current state only
- ▶ ergodic chains converge to unique distribution ( $t \rightarrow \infty$ )
- ▶ Foster-Lyapunov functions for bounds

# Markovian population models

## Approaches to their analysis

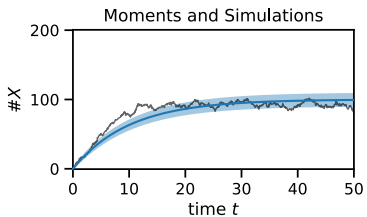


### Moment equation

Moments such as mean  $E(X_t)$  and variance  $E(X_t^2) - E(X_t)^2$  are described by (often linear) ODEs.

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

## Moment-based Methods



# Markovian population models

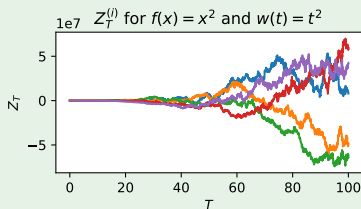
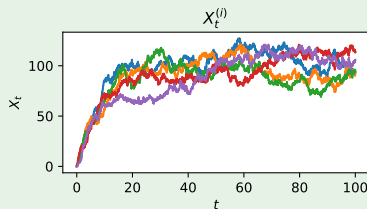
## Martingale process

- ▶ start with the moment ODE
- ▶ multiply time-weighting:  $w(t) = t^k$ ,  $k \in \mathbb{N}$  or  $w(t) = \exp(\lambda t)$
- ▶ analytic integration results in a martingale

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

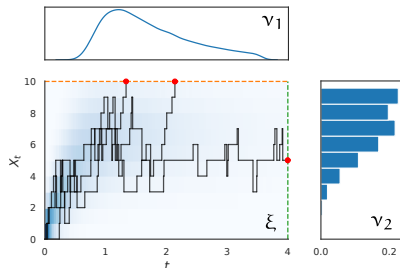
- ▶ known expectation:  $E(Z_T) = 0$ ,  $\forall T \geq 0$

## Example



# Bounding mean first-passage times

## Linear moment constraints



- ▶  $\tau = \inf\{X_t \geq H \mid t \geq 0\} \wedge T$
- ▶ time weighting  $w(t) = t^k$
- ▶ exp. occupation measure  $\xi$
- ▶ exit location measures  $\nu_1, \nu_2$

## Linear moment constraint

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



# Bounding mean first-passage times

## Moment matrices

### Moment matrices

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

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

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

### Example

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

[BBW20]

# Bounding mean first-passage times

## Semi-definite program

- ▶ 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

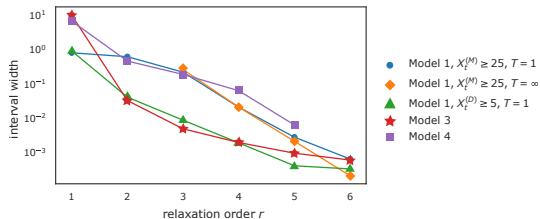
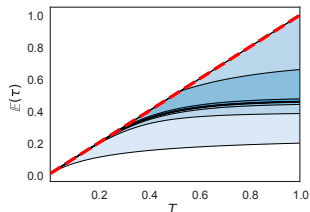
1. *linear constraints* on moments and
  2. *positive semi-definite constraints* on certain matrices.
- ▶ alternative: linear *Hausdorff constraints* instead of semi-definite constraints

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

[BBW20]

# Bounding mean first-passage times

## Results

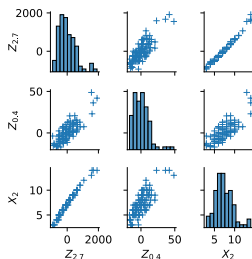


- ▶ fast convergence of bounds with increasing order
- ▶ SDPs are usually solved within seconds
- ▶ numerically challenging (inherent stiffness)
- ▶ scaling state-space / model size is difficult

[BBW20]

# Linear control variates

Using correlated RVs with known expected value



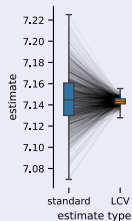
- ▶ improve MC estimates using  $Z_T$
- ▶ use correlations between  $Z_T$  and  $X_T$
- ▶  $E(X_T + bZ_T)$  instead of  $E(X_T)$   
(recall  $E(Z_T) = 0$ )
- ▶ time-weighting  $w(t) = \exp(\lambda t)$

## Linear control variates

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

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

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



# Linear control variates

## Efficiency trade-off

- ▶ infinite possible  $Z$
- ▶ different time-weighting  $\lambda \rightarrow$  different correlation
- ▶ the trade-off:

### cost: slowdown

$$\text{cost}_{\text{old}} / \text{cost}_{\text{new}}$$

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

### benefit: variance reduction

$$\sigma_{\text{new}}^2 / \sigma_{\text{old}}^2$$

- ▶ highly correlated variates

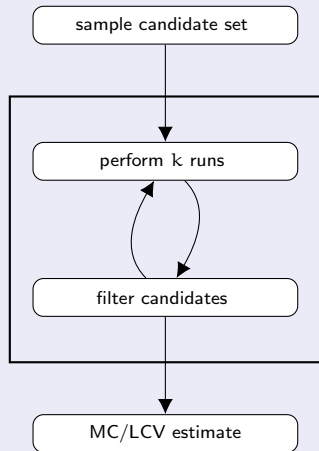
- ▶ approach: correlations between candidates and the target RV

[BBW19; BBW22]

# Linear control variates

Selection by filtering

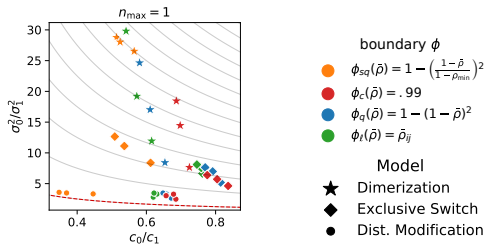
## Algorithm 1



[BBW19]

filter criteria:

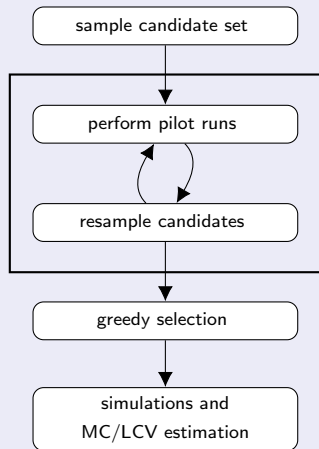
1. low target correlation
2. various redundancy heuristics



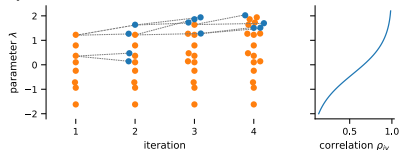
# Linear control variates

## Selection by resampling

### Algorithm 2



- ▶ *resampling*: proportional to improvement

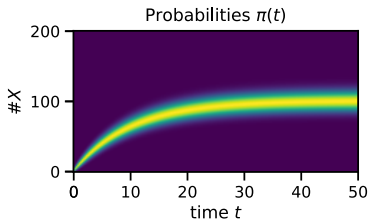


- ▶ *selection*: greedy by improvement

Results:

- ▶ performance equal/better than Alg. 1
- ▶ less hyper-parameter headaches

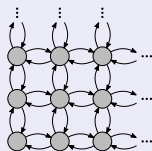
## Aggregation & refinement





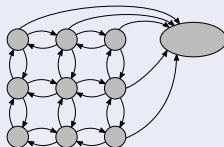
# State-space truncation

## Original



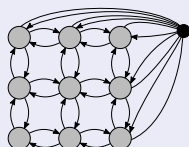
- ▶ very large/infinite
- ▶ impossible to analyze

## Sink state



- ▶ transient analysis
- ▶ keep track of approx. error

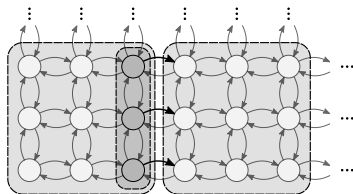
## Redirection



- ▶ stationary dist.
- ▶ dependent on redirection

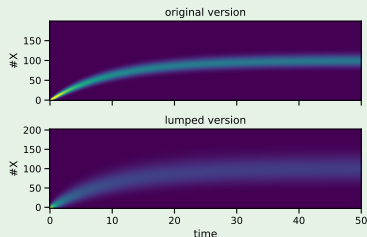
# State-space aggregation

Treating hyper-cubes of states as one



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

## Example (birth-death process)



- ▶ resulting distribution more “flat”
- ▶ main probability masses coincide

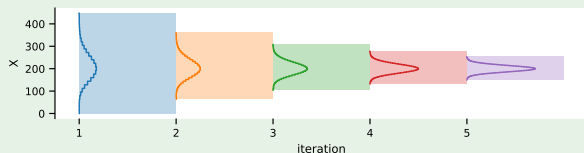
# Stationary distribution

## 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

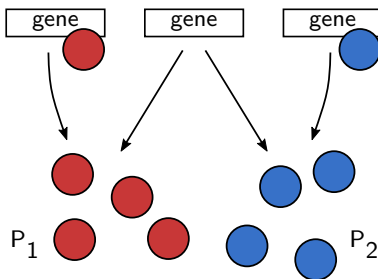
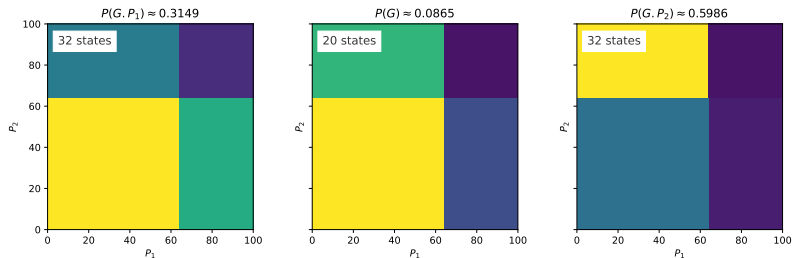
### Example (birth-death process)



[Bac+21b; Bac+21a]

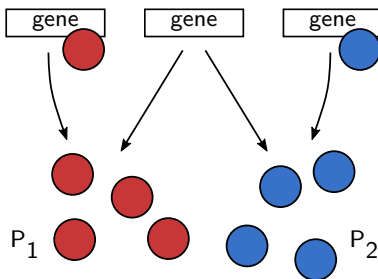
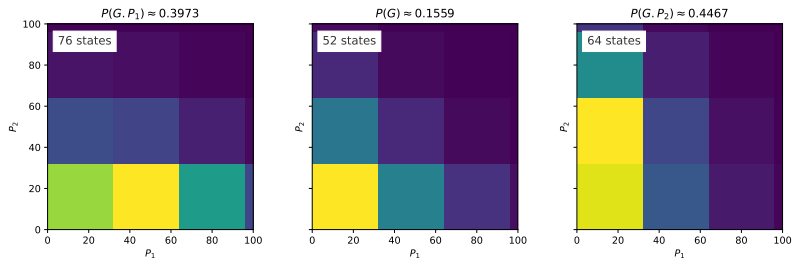
# Stationary distribution

## Example



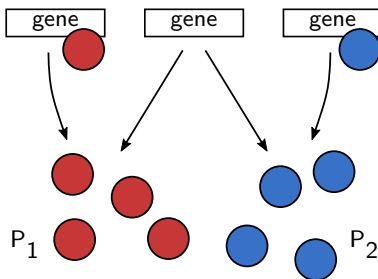
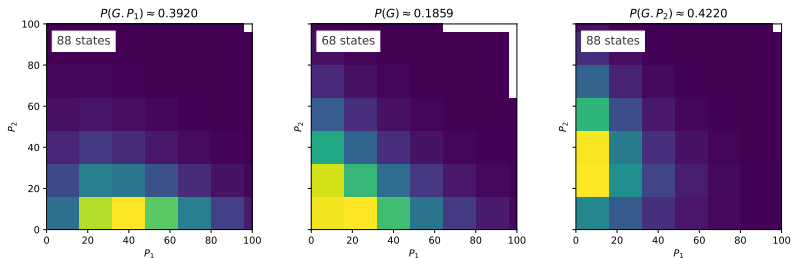
# Stationary distribution

## Example



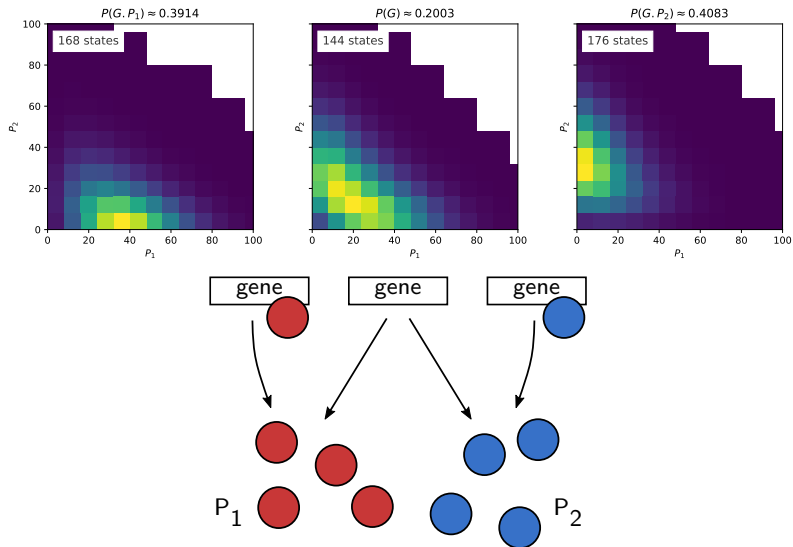
# Stationary distribution

## Example



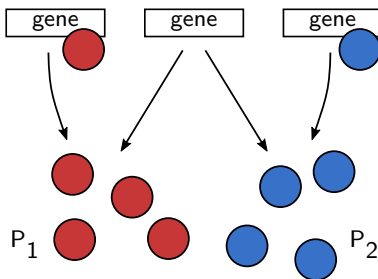
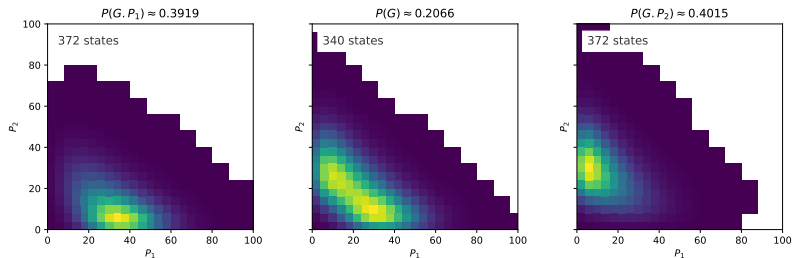
# Stationary distribution

## Example



# Stationary distribution

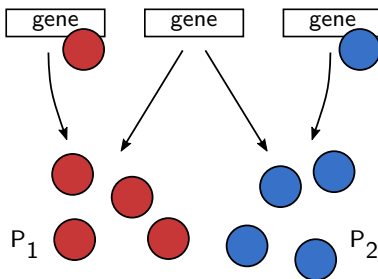
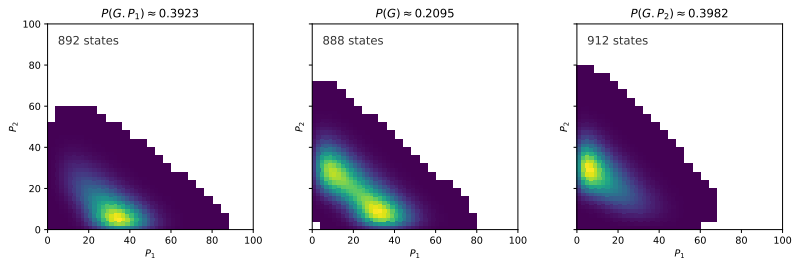
## Example





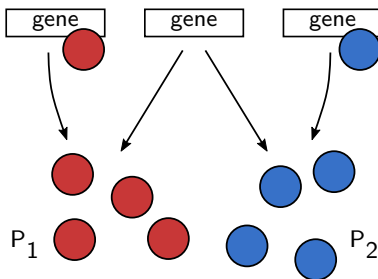
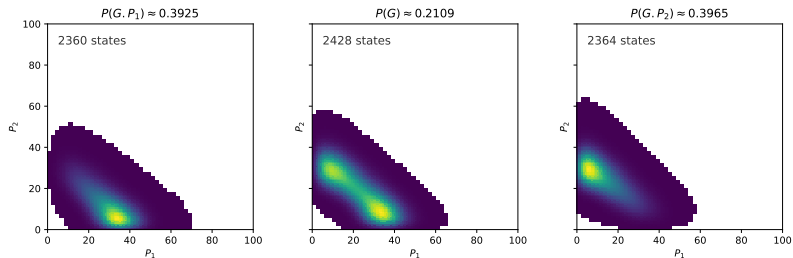
# Stationary distribution

## Example



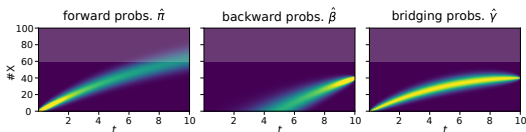
# Stationary distribution

## Example



# Bridging problem

Dynamical analysis under initial *and* terminal constraints



## Forward probabilities $\pi$

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

## Backward probabilities $\beta$

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

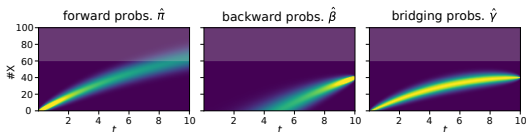
## Bridging probabilities $\gamma$

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

[Bac+21b]

# Bridging problem

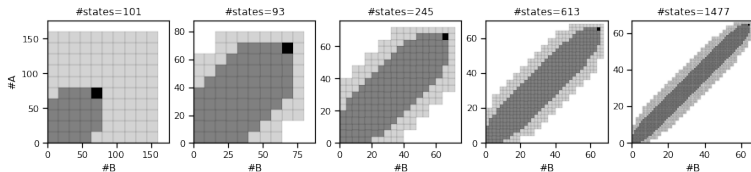
## Refinement



- 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  $\hat{\gamma}(x_i, t)$



# Bridging problem

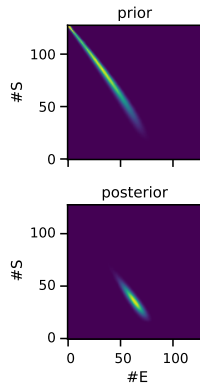
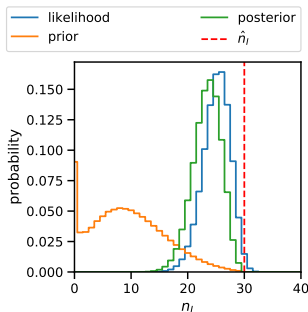
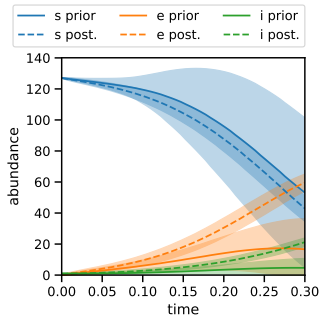
## Bayesian filtering in an SEIR model



We know:

- ▶ initial state
- ▶ uncertain measurement of I at  $T = 0.3$

We are interested in the posterior at T.

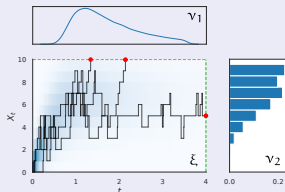


[Bac+21b]

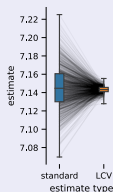
# Contributions

## Moment-based

- ▶ bounding of mean first-passage times [BBW20]



- ▶ variance reduction for MC estimation [BBW19; BBW22]

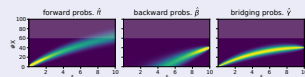


## Aggregation & refinement

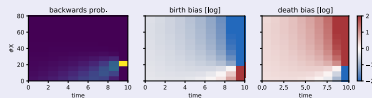
- ▶ stationary distribution [Bac+21a]



- ▶ bridging distribution [Bac+21b]



- ▶ importance sampling



## Augmented Lyapunov functions

local alteration of valid Lyapunov functions for tighter guarantees

## 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.

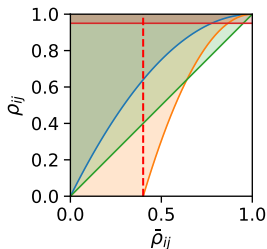
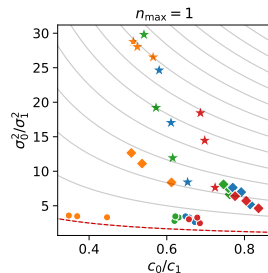
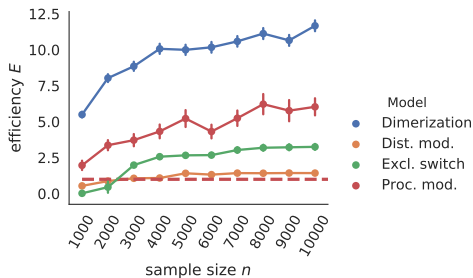
## References II

- [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.



# Linear control variates

## Results



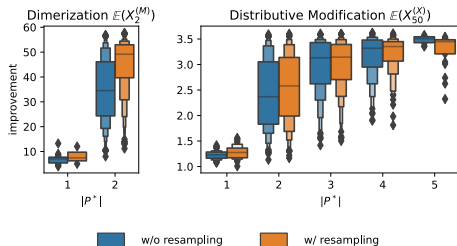
- boundary  $\phi$
- orange circle:  $\phi_{sq}(\bar{\rho}) = 1 - \left(\frac{1-\bar{\rho}}{1-\bar{\rho}_{\min}}\right)^2$
  - red circle:  $\phi_c(\bar{\rho}) = .99$
  - blue circle:  $\phi_q(\bar{\rho}) = 1 - (1-\bar{\rho})^2$
  - green circle:  $\phi_t(\bar{\rho}) = \bar{\rho}_{ij}$

- boundary  $\phi$
- orange circle:  $\phi_{sq}(\bar{\rho}) = 1 - \left(\frac{1-\bar{\rho}}{1-\bar{\rho}_{\min}}\right)^2$
  - red circle:  $\phi_c(\bar{\rho}) = .99$
  - blue circle:  $\phi_q(\bar{\rho}) = 1 - (1-\bar{\rho})^2$
  - green circle:  $\phi_t(\bar{\rho}) = \bar{\rho}_{ij}$
- Model
- ★ Dimerization
  - ◆ Exclusive Switch
  - Dist. Modification

# Linear control variates

## Results

- ▶ SMC can improve variance reduction



- ▶ less dependence on initial covariates
- ▶ more consistent performance