Metastability in Biological Systems

Susanna Röblitz

Computational Biology Unit
Department of Informatics
University of Bergen

Woudschoten Conference, September 25, 2024


NORWEGIAN SEA
Bergen
NORWAY
Oslo
SWEDEN
DENMARK
© 2010 EB, Inc.
World Heritage Site

Outline

Biological background:

- the central dogma of molecular biology
- stochastic gene expression

Mathematical modelling:

- stochastic chemical kinetics and the chemical master equation
- Gillespie's stochastic simulation algorithm (SSA)

Markov state modelling (MSM):

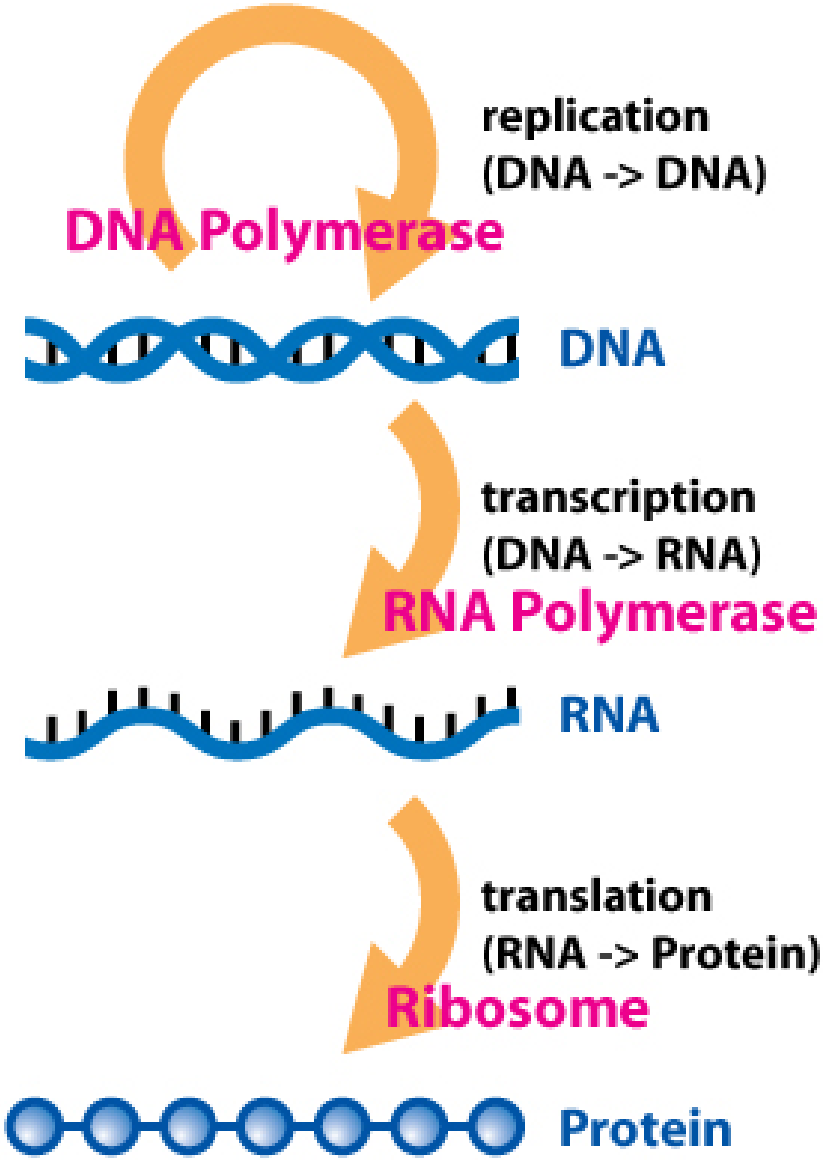
- metastability
- robust Perron cluster analysis (PCCA+)
- discretization and error estimation

Summary:

- conclusion
- future work

Github repository with examples: `https://github.com/sroeblitz/stochGRN`

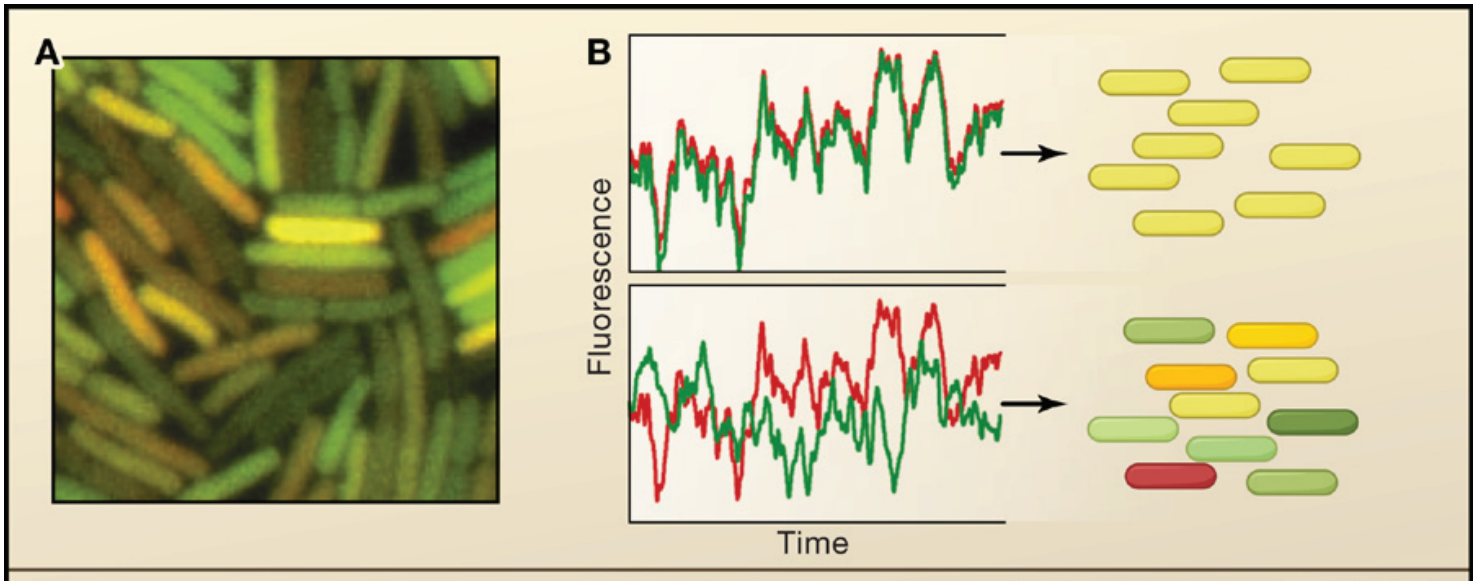
The central dogma of molecular biology



- genes encode proteins and proteins dictate cell function
- potential control points for self-regulation by adjusting the amount and type of proteins

[Dhorspool at en.wikipedia, CC BY-SA 3.0, via Wikimedia Commons]

Stochastic gene expression



Fluorescence imaging of individual E. coli [Raj et al. 2008. DOI 10.1016/j.cell.2008.09.050]

Imaging technologies reveal marked variability in protein expression due to **extrinsic** and **intrinsic** noise.

Stochastic modelling

Cells are inherintly noisy biochemical reactors.

Extrinsic noise:

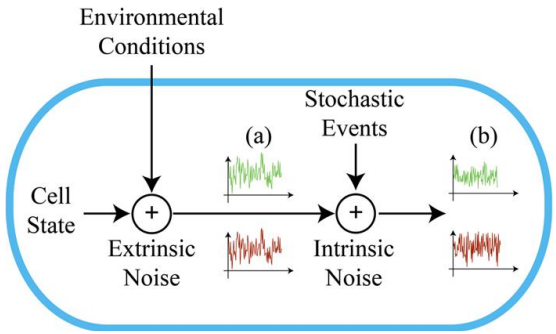
- different cell cycle stages
- spatial variation in environmental signals across a population of cells
- cell-to-cell differences in energy budget (particularly ATP levels)
- random partitioning of molecules at cell division
- ...

→ model parameters and initial values are random variables

Extrinsic noise can affect levels and types of **intrinsic noise**:

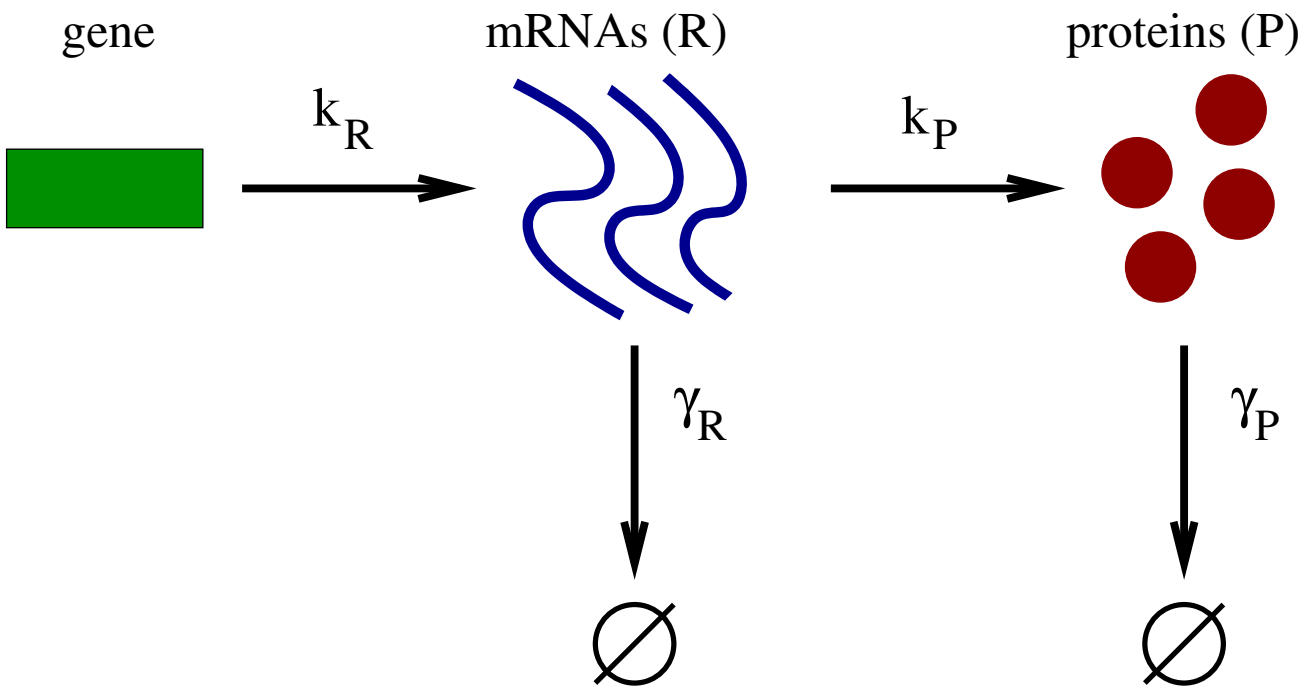
- random collisions between reactants due to low copy-number effects (including discrete birth and death events) and diffusive dynamics

→ chemical reactions are modelled as a Markov jump process (CME)



Modelling stochastic gene expression

Gene expression can be modeled as a systems of coupled stochastic reactions.



Stochastic chemical kinetics

General reaction scheme (D species, R reactions):

$$\sum_{i=1}^{D} m_{ir} x_i \xrightarrow{\alpha_r(x)} \sum_{i=1}^{D} n_{ir} x_i, \quad r = 1, \ldots, R$$

$x = (x_1, \ldots, x_D) \in \mathbb{N}^D$: state of the system (molecule numbers of species)

$m_{ir}, n_{ir} \in \mathbb{N}$: stoichiometric coefficients

$\alpha_r(x) : \mathbb{N}^D \to \mathbb{R}_{\geq 0}$: reaction propensity, $r = 1, \ldots, R$

$$x \xrightarrow{\alpha_r(x)} x + \nu_r(x), \quad \nu_r = \begin{pmatrix} n_{1r} - m_{1r} \\ n_{2r} - m_{2r} \\ \vdots \\ n_{Nr} - m_{Nr} \end{pmatrix} \in \mathbb{Z}^D$$

The probability of reaction r taking place in the infinitesimal time interval $[t, t + dt)$ is given by $\alpha_r(x(t))dt$.

The chemical master equation (CME)

Assumption: dt is so small that at most one reaction can take place over $[t, t+dt)$.

$$p(x, t+dt) = \left(1 - \sum_{r=1}^{R} \alpha_r(x)dt\right) p(x,t) + \sum_{r=1}^{R} \alpha_r(x-\nu_r)dtp(x-\nu_r, t)$$

$$\frac{p(x,t+dt) - p(x,t)}{dt} = \sum_{r=1}^{R} (\alpha_r(x-\nu_r)p(x-\nu_r,t) - \alpha_r(x)p(x,t))$$

$$\boxed{\frac{dp(x,t)}{dt} = \sum_{r=1}^{R} (\alpha_r(x-\nu_r)p(x-\nu_r,t) - \alpha_r(x)p(x,t))}$$

linear ODE system with one ODE for each possible state
$\Rightarrow$ compute single realizations rather than entire distribution

The stochastic simulation algorithm (SSA)

$P_0(\tau|x,t)$: probability that no reaction takes place in the time interval $[t, t+\tau)$, given $X(t) = x$

Assumption: what happens over $[t, t+\tau)$ is independent of what happens over $[t+\tau, t+\tau+d\tau)$ (Markov property)

$$\underbrace{P_0(\tau + d\tau|x,t)}_{\text{no reaction over}[t,t+\tau+d\tau)} = \underbrace{P_0(\tau|x,t)}_{\text{no reaction over}[t,t+\tau)} \cdot \underbrace{\left(1 - \sum_{j=1}^{R} \alpha_j(x) d\tau\right)}_{\text{no reaction over}[t+\tau,t+\tau+d\tau)}$$

$$\frac{P_0(\tau + d\tau|x,t) - P_0(\tau|x,t)}{d\tau} = -\alpha_{sum}(x) P_0(\tau|x,t), \quad \alpha_{sum}(x) := \sum_{j=1}^{R} \alpha_j(x)$$

$d\tau \to 0$: linear scalar ODE with $P_0(0|x,t) = 1$ and solution

$$\boxed{P_0(\tau|x,t) = \exp(-\alpha_{sum}(x)\tau)}$$

Stochastic simulation algorithm (SSA)

Key quantity for SSA:
$p(\tau, j|x, t)d\tau$: probability that the next reaction will (a) be reaction j and (b) occur in the time interval $[t+\tau, t+\tau+d\tau)$ given $X(t) = x$

$$p(\tau, j|x, t)d\tau = \underbrace{P_0(\tau|x,t)}_{\text{no reaction over}[t,t+\tau)} \cdot \underbrace{\alpha_j(x)d\tau}_{\text{reaction j took place over}[t+\tau,t+\tau+d\tau)}$$

$$p(\tau, j|x, t) = \underbrace{\frac{\alpha_j(x)}{\alpha_{sum}(x)}}_{\text{next reaction index}} \cdot \underbrace{\alpha_{sum}(x)\exp(-\alpha_{sum}(x)\tau)}_{\text{time until next reaction}}$$

$\rightarrow$ independent sampling of **reaction index** (chance of picking reaction j is proportional to $\alpha_j(x)$) and **reaction time** (exponentially distributed) via uniform $(0, 1)$ sample

Stochastic Simulation Algorithm

Sample trajectories of that process can be generated by the Stochastic Simulation Algorithm (SSA) [D. T. Gillespie (1976). https://doi.org/10.1016/0021-9991(76)90041-3]

1. Set $t = 0$ and assign the initial number of molecules $X(0)$.
2. Draw two uniform random numbers u_1 and u_2 in $(0, 1)$.
3. Compute the total reaction intensity $\alpha_{sum}(X(t))$. Generate the *time to the next reaction* τ by setting

$$\tau := -\log u_1 / \alpha_{sum}(X(t)).$$

   Determine the *next reaction* j by the requirement that

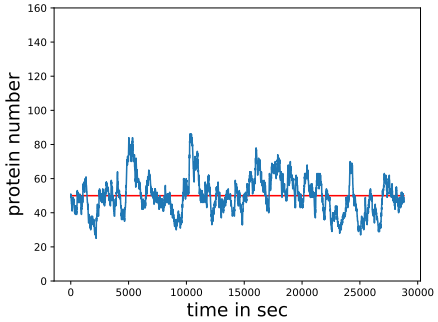
$$\sum_{s=1}^{j-1} \alpha_s(X(t)) < \alpha_{sum}(X(t)) u_2 \leq \sum_{s=1}^{j} \alpha_s(X(t)).$$

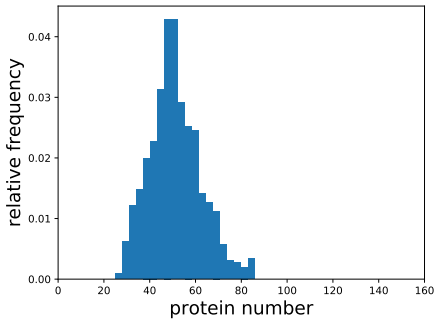
4. Update $t := t + \tau$ and $X(t + \tau) := X(t) + \nu_j$.
5. Repeat from 1. until some final time T is reached.

Simulating stochastic gene expression

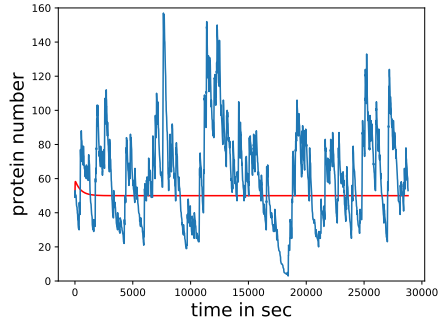
Noise in **prokaryotic** *gene expression depends on the rates of transcription and translation.* [Ozbudak et al. (2002). Regulation of noise in the expression of a single gene. DOI: 10.1038/ng869]

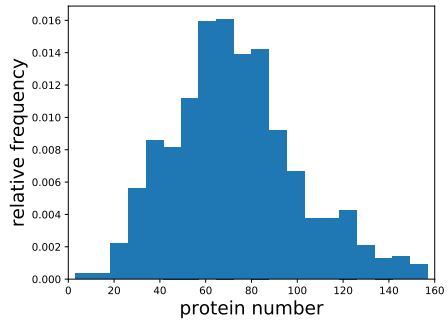
high transcription, low translation





low transcription, high translation





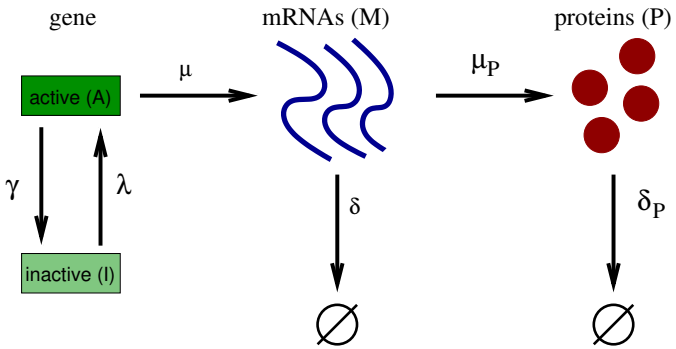
bursts of protein creation of average size $b = k_P/\gamma_R$ occurring at average rate k_R

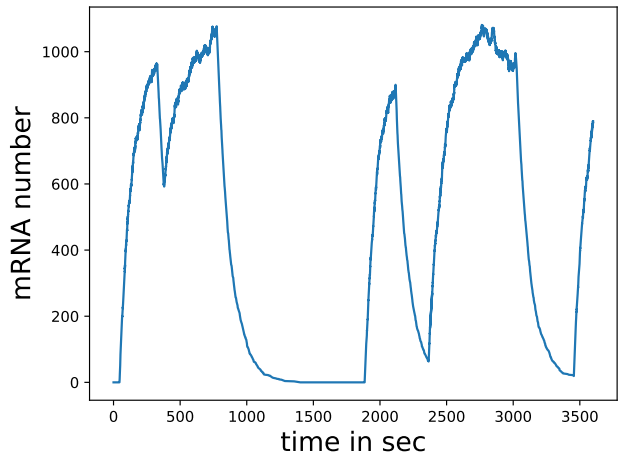
Code: `stoch-gene-expression.py`

Transcriptional regulation

Gene expression is more complicated in **eukaryotic** cells!

[Raj et al. (2006). Stochastic mRNA Synthesis in Mammalian Cells. DOI: 10.1371/journal.pbio.0040309]



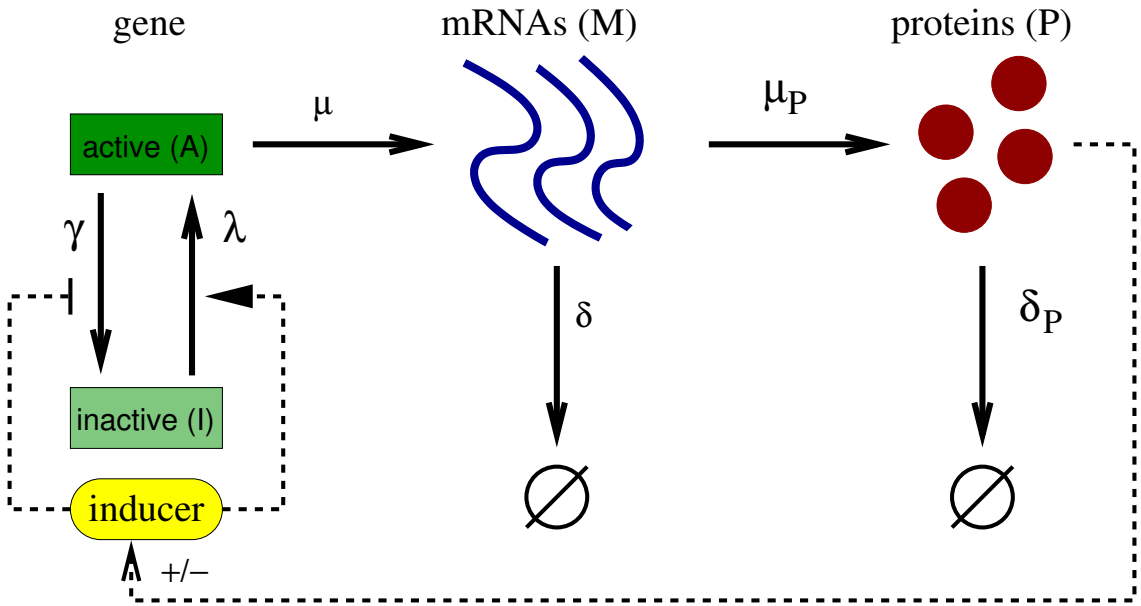


- transcripts are modified in the nucleus before they are exported to the cytoplasm for translation (transcriptional regulation)
- A stochastic model of gene activation and inactivation can explain transcriptional bursting
- mRNA expression can be buffered at the protein level by slow protein degradation rates

Code:

`stoch-gene-expresssion-RNA-bursts.py`

Transcriptional regulation

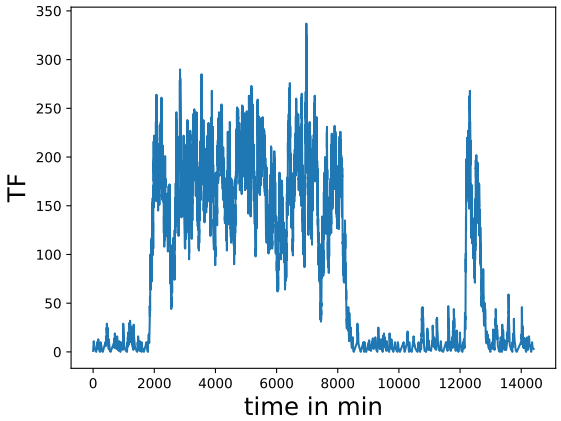


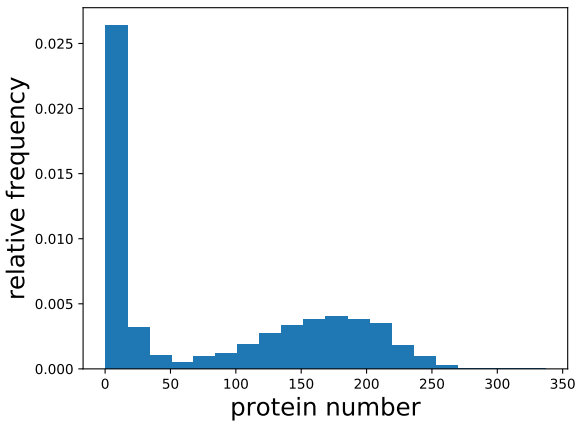
- rate of switching between active and inactive transcription state can depend upon an external **inducer** (a molecule that regulates gene expression by disabling repressors or binding to activators)
- a protein can modulate the expression of its own gene (**auto-regulation**)

Bistability

Simple model of auto-activation reveals bistability

[Hermsen et al. (2011). Speed, Sensitivity, and Bistability in Auto-activating Signaling Circuits. DOI: 10.1371/journal.pcbi.1002265]





$$c'(t) = g(rc)b/V - \beta c$$

$$g(rc) = \alpha \frac{(rc/K)^H + 1/f}{(rc/K)^H + 1}$$

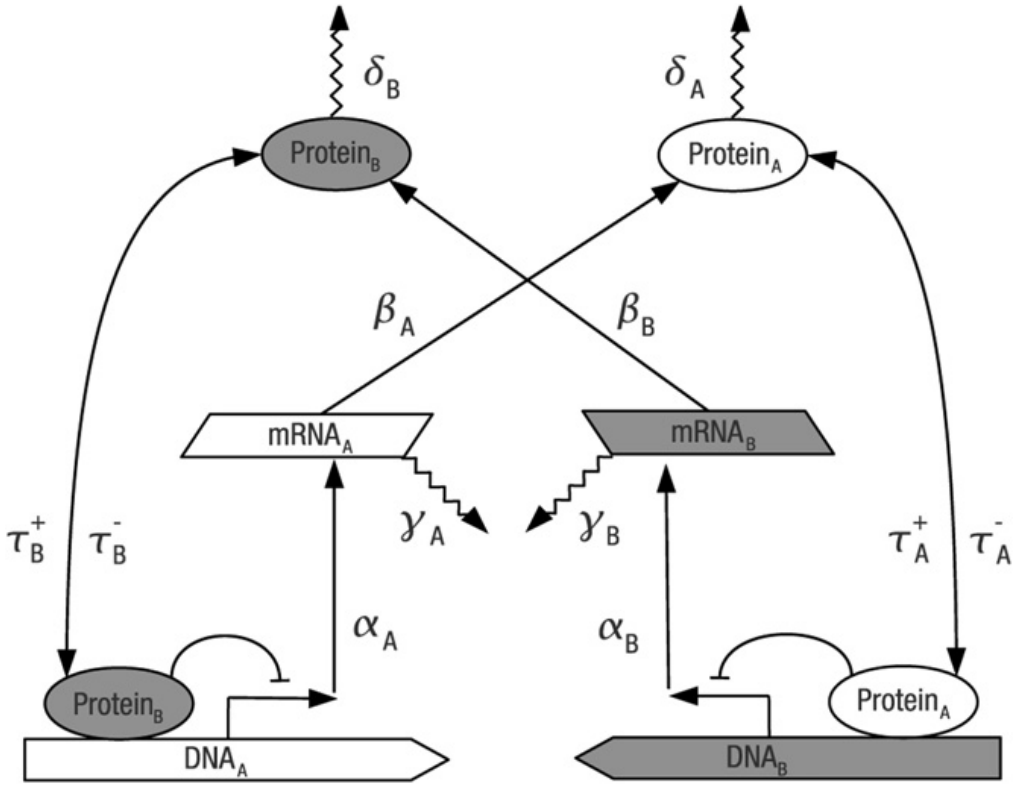
c: TF concentration
r: fraction of activated transcription factors (TF)
b: burst size (each mRNA transcribed from the promoter is instantly translated b times)
V: volume of the cell
β: degradation rate constant of the TF
α: maximal transcription rate at full activation
K: dissociation constant of the modified TF binding to its operator
f: maximal fold change of the promoter ($> \alpha/\beta$ for bimodality)
H: Hill coefficient

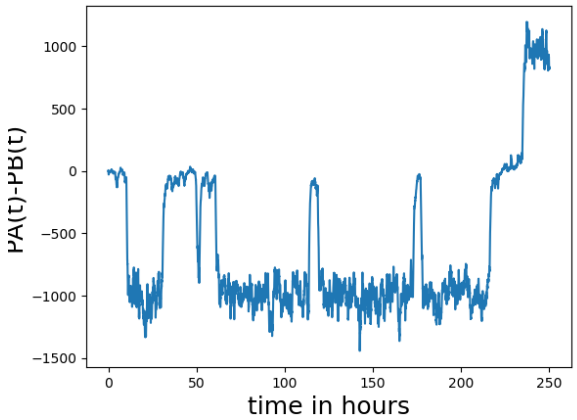
Code: `stoch-gene-expression-1D-bistable.py`

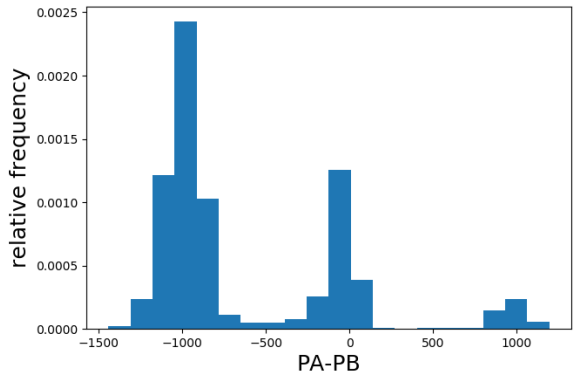
Multistability

A network consisting of two mutually inhibiting genes displays multistability.

[Strasser et al. (2012). Stability and multiattractor dynamics of a toggle switch based on a two-stage model of stochastic gene expression. doi: 10.1016/j.bpj.2011.11.4000]





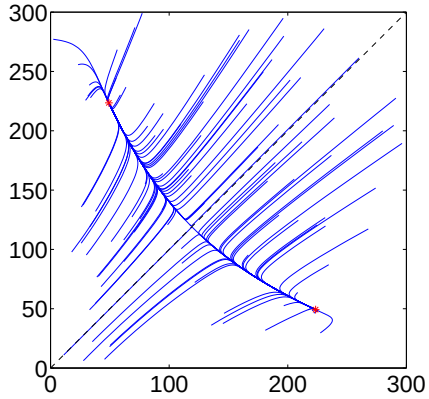


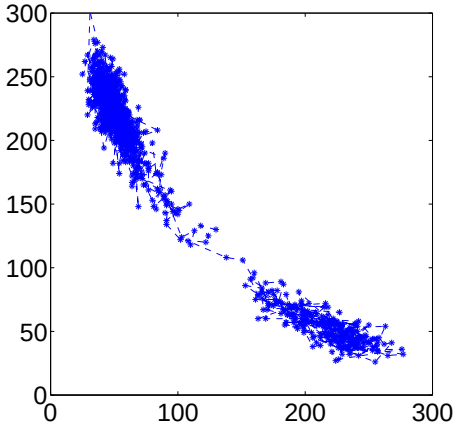
Code: `stoch-gene-expression-2D-multistable.py`

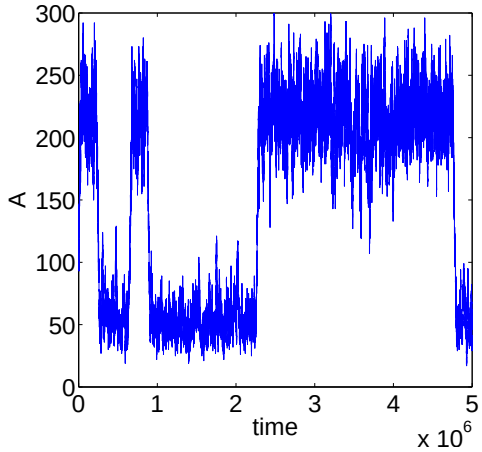
Rare events

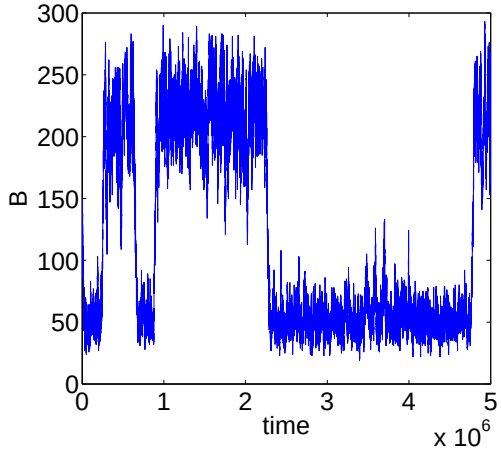
Semi-mechanistic model of the genetic toggle switch [Gardner et al., Nature 403 (2000)]

reaction	mechanism	propensity
r_1	$\star \to A$	$\alpha_1 = c_1/(c_2 + B^\beta)$
r_2	$A \to \star$	$\alpha_2 = c_3 A$
r_3	$\star \to B$	$\alpha_3 = c_4/(c_5 + A^\gamma)$
r_4	$B \to \star$	$\alpha_4 = c_6 B$



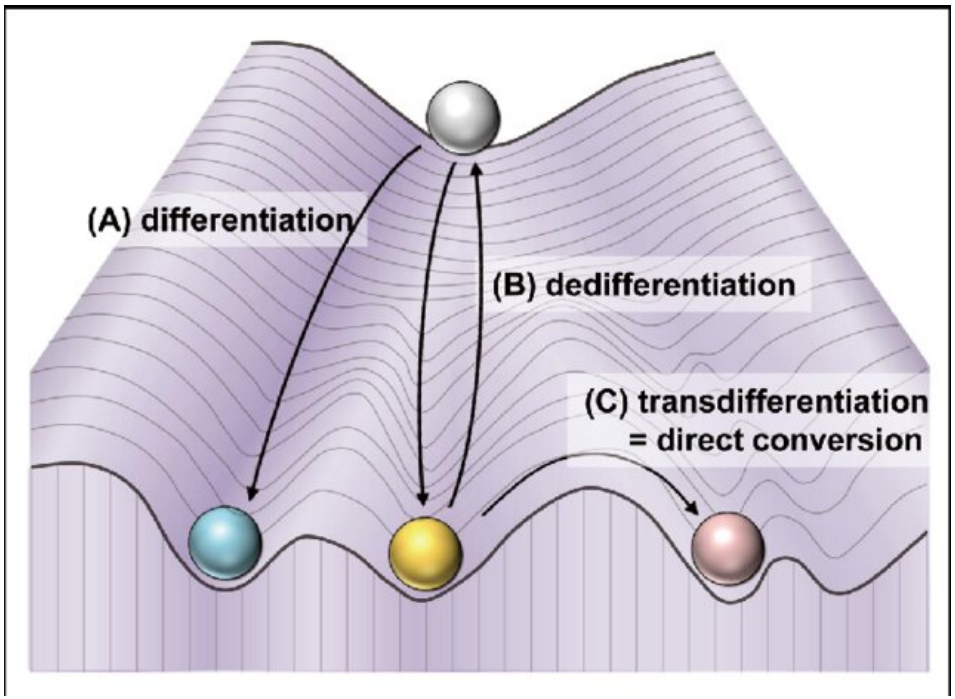






only 5 transitions between $\{(A,B) : A > B\}$ and $\{(A,B) : A < B\}$ within $5 \cdot 10^6$ time steps $\Rightarrow$ poor statistics: 3:2 (theoretically 1:1)

Gene-regulatory networks

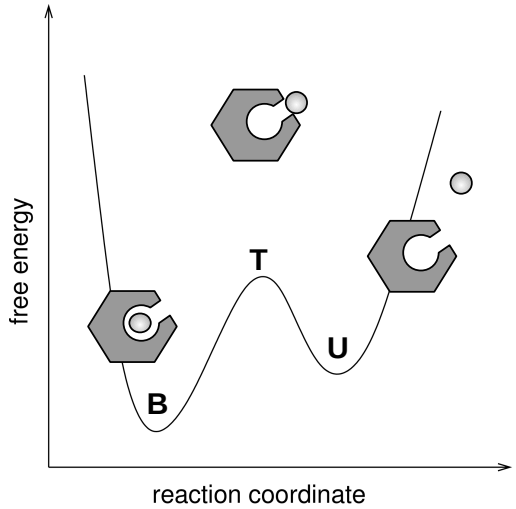


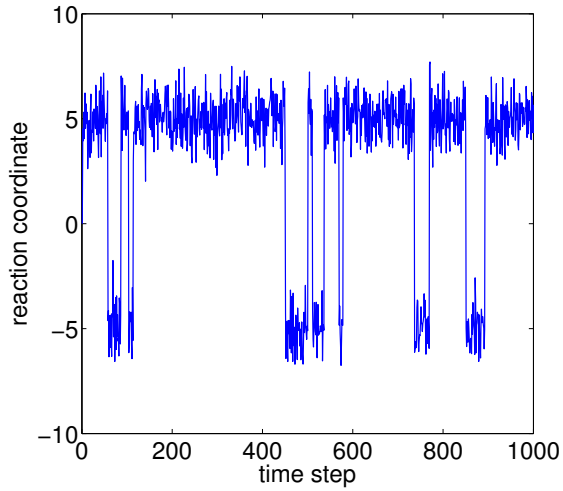
Waddington's epigenetic landscape
[Han et al. (2022). https://doi.org/10.1253/circj.CJ-21-0703]

- probabilistic framework results in **complex multi-attractor dynamics**
- attractors can be identified with committed and primed states in cell differentiation (**cellular phenotypes**)
- since the dynamics can switch between different attractors, we call it **metastable** instead of multistable

How can we characterize the dynamics of these expression patterns?

Detour: Molecular conformation dynamics





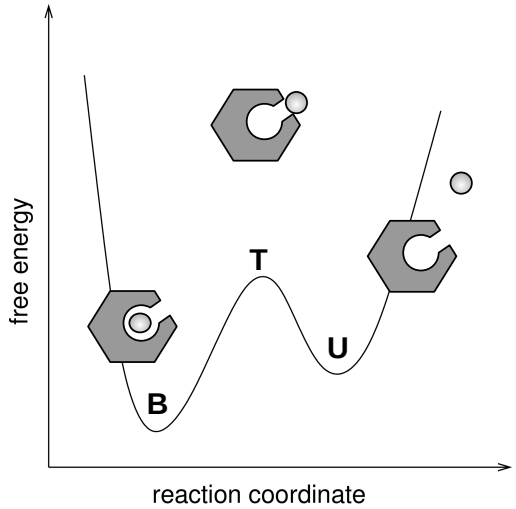
Identification of molecular conformations through the dominant eigenfunctions of the transfer operator $\mathcal{T}^\tau f(q) = \int_{\mathbb{R}^d} f(\Pi_q \Phi^{-\tau}(q,p))\eta(p)\,dp$

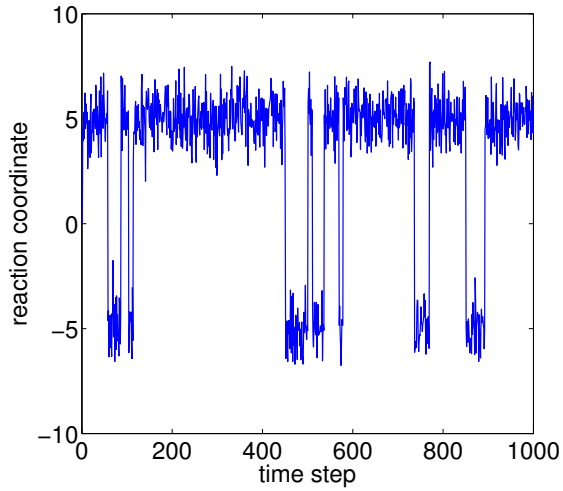
[Schütte, Fischer, Huisinga, Deuflhard (1999). https://doi.org/10.1006/jcph.1999.6231]

Molecular Dynamics (MD)

- continuous state space
- self-adjoint transfer operator (reversible dynamics)
- Boltzmann distribution (importance sampling!)

Detour: Molecular conformation dynamics





Identification of molecular conformations through the dominant eigenfunctions of the transfer operator $\mathcal{T}^\tau f(q) = \int_{\mathbb{R}^d} f(\Pi_q \Phi^{-\tau}(q,p))\eta(p)\,dp$

[Schütte, Fischer, Huisinga, Deuflhard (1999). https://doi.org/10.1006/jcph.1999.6231]

Molecular Dynamics (MD)

- continuous state space
- self-adjoint transfer operator (reversible dynamics)
- Boltzmann distribution (importance sampling!)

Gene-regulatory networks (GRN)

- discrete state space
- transfer operator not self-adjoint
- unknown stationary distribution

Metastability

$$\partial_t p(x,t) = \sum_{r=1}^{R} [\alpha_r(x-\nu_r)p(x-\nu_r,t) - \alpha_r(x)p(x,t)] =: \mathcal{M}p(x,t),$$

Adjoint operator: $\mathcal{M}^* q = \sum_{r=1}^{R} \alpha_r(x)[q(x+\nu_r) - q(x)]$

Conservation of mass:

$$\sum_{x\in\mathbb{N}^D} \mathcal{M}p(x,t) = 0$$

Metastability

$$\partial_t p(x,t) = \sum_{r=1}^{R} [\alpha_r(x-\nu_r)p(x-\nu_r,t) - \alpha_r(x)p(x,t)] =: \mathcal{M}p(x,t),$$

Adjoint operator: $\mathcal{M}^* q = \sum_{r=1}^{R} \alpha_r(x)[q(x+\nu_r) - q(x)]$

Conservation of mass:

$$\sum_{x\in\mathbb{N}^D} \mathcal{M}p(x,t) = 0$$

Idea

For a **metastable** set $\Omega \in \mathbb{N}^D$, i.e. an area where a trajectory stays for a long time before it switches to another metastable set, this conservation of mass should still hold approximately:

$$\sum_{x\in\Omega} \mathcal{M}p(x,t) \approx 0$$

From sets to functions

Definition

We call a set of functions $\{C_k : \mathbb{N}^D \to [0,1]\}$ with $\sum_k C_k(x) = \mathbb{1}_{\mathbb{N}^D} \, \forall \, x \in \mathbb{N}^D$ a **metastable function partitioning** if for any probability distribution $p(x,t)$ and all functions $C_k(x)$:

$$\sum_{x \in \mathbb{N}^D} C_k(x) \mathcal{M} p(x,t) \approx 0$$

From sets to functions

Definition

We call a set of functions $\{C_k : \mathbb{N}^D \to [0,1]\}$ with $\sum_k C_k(x) = \mathbb{1}_{\mathbb{N}^D} \, \forall \, x \in \mathbb{N}^D$ a **metastable function partitioning** if for any probability distribution $p(x,t)$ and all functions $C_k(x)$:

$$\sum_{x \in \mathbb{N}^D} C_k(x) \mathcal{M} p(x,t) \approx 0$$

This definition can be restated in terms of the adjoint $\mathcal{M}^*$ as

$$\mathcal{M}^* C_k(x) \approx 0$$

Eigenvalue problem

Metastable functions can be identified as right eigenfunctions of the adjoint operator $\mathcal{M}^*$ for eigenvalues close to zero:

$$\mathcal{M}^* C_k(x) = \lambda_k C_k(x), \quad \lambda_k \approx 0$$

Finite state space

Finite-dimensional subspace $\Omega \subset \mathbb{N}^D$: the CME operator can be written as reaction rate matrix

$$M_{ij} = \begin{cases} -\sum_{r=1}^{R} \alpha_r(x_i), & \text{for } i = j \\ \alpha_r(x_i), & \text{for all } j \text{ such that } x_j = x_i + \nu_r \\ 0, & \text{otherwise} \end{cases}$$

In matrix notation, the CME reads

$$\dot{\mathbf{p}}^\top = \mathbf{p}^\top M$$

where $\mathbf{p} = [p(x_1), p(x_2), \ldots]^\top$.

If M is neither decomposable nor of splitting type, M has a unique stationary distribution $\boldsymbol{\pi} = [\pi(x_1), \pi(x_2), \ldots]^\top$ satisfying

$$\boldsymbol{\pi}^\top M = 0.$$

Discrete eigenvalue problem:

$$MX = X\Lambda, \quad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n_C}),\ \lambda_1 = 0,\ \lambda_{2,\ldots,n_c} < 0$$

From eigenvectors to membership vectors

Aim: Transform the invariant subspace X to membership vectors χ,

$$\chi = XA \quad (A \text{ regular}) \quad \text{s.th.}$$

1. $\sum_{J=1}^{n_c} \chi_J(i) = 1 \quad \forall i \in \{1, \ldots, N\}$ (partition of unity)
2. $\chi_J(i) \geq 0 \quad \forall i \in \{1, \ldots, N\},\ J \in \{1, \ldots, n_C\}$ (positivity)

How can we find this transformation A?

From eigenvectors to membership vectors

Aim: Transform the invariant subspace X to membership vectors χ,

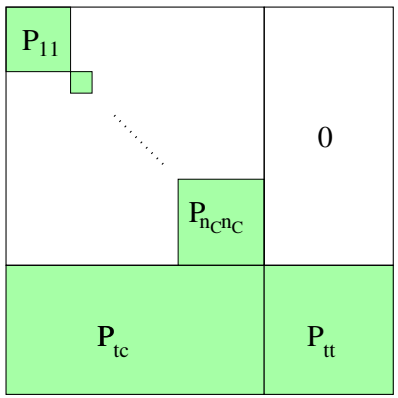
$$\chi = XA \quad (A \text{ regular}) \quad \text{s.th.}$$

1. $\sum_{J=1}^{n_c} \chi_J(i) = 1 \quad \forall i \in \{1, \ldots, N\}$ (partition of unity)
2. $\chi_J(i) \geq 0 \quad \forall i \in \{1, \ldots, N\},\ J \in \{1, \ldots, n_C\}$ (positivity)

How can we find this transformation A?

The transition probability matrix $P(\tau) := \exp(\tau \cdot M)$ has the same eigenvectors!

The transition probability matrices of metastable Markov jump processes (and their eigenvectors) have (upon reordering of states) a **characteristic structure**:



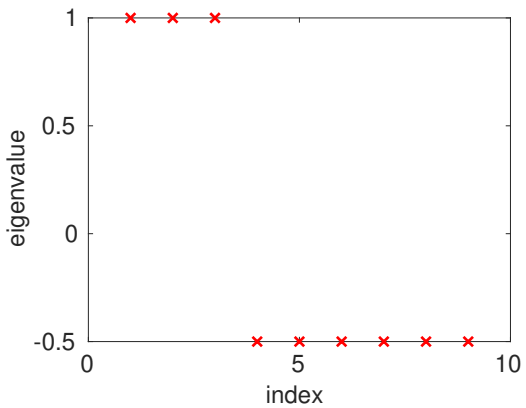
- high transition probabilities between the states of a metastable set (blocks)
- small transition probabilities to states outside a metastable set
- transition states

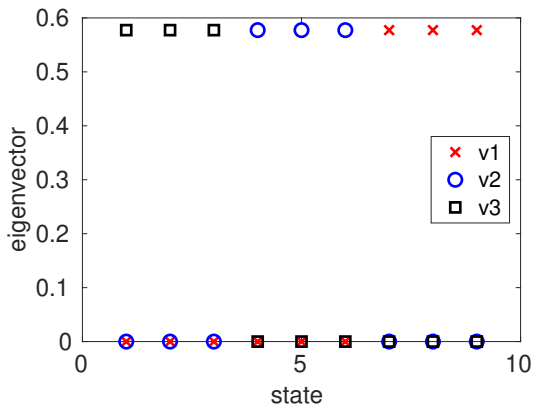
Decoupled Markov chains

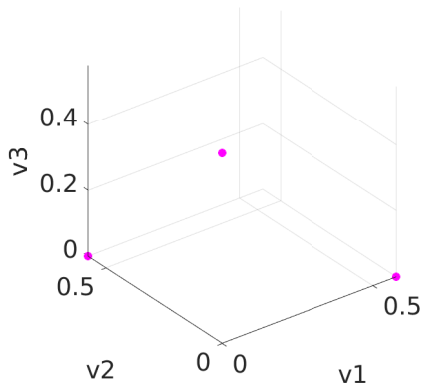
The dominant eigenvectors are constant on the blocks!

Example:

$$P = \begin{pmatrix} 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 \end{pmatrix}$$



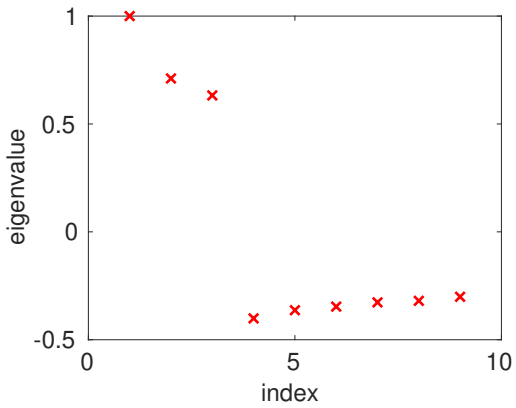




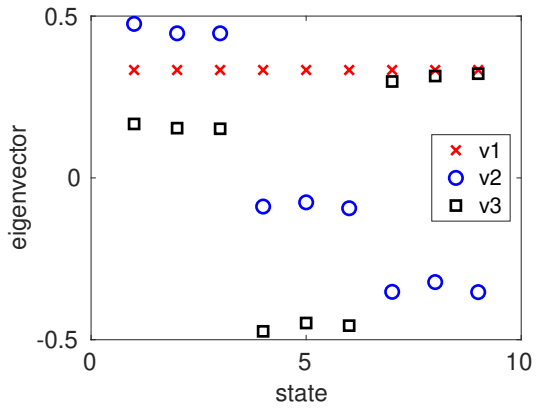
Nearly decoupled Markov chains

This structure is very robust upon perturbations!

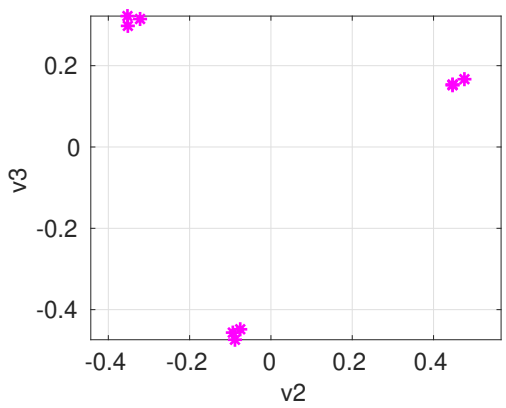
```
eps=1e-1;
P=P+eps*rand(9,9);
P=0.5*(P+P');   %reversibility
P=diag(1./sum(P,2))*P;  %stochasticity
```



eigenvalues



eigenvectors



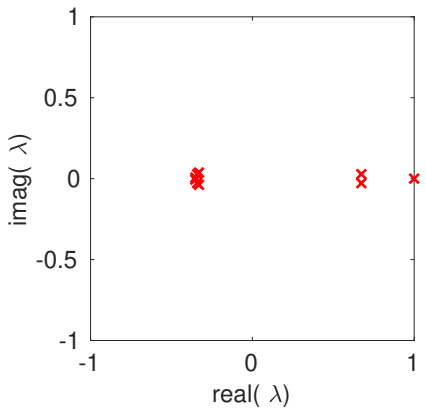
simplex structure

[P. Deuflhard, M. Weber (2005). Robust Perron cluster analysis in conformation dynamics. doi: 10.1016/j.laa.2004.10.026]

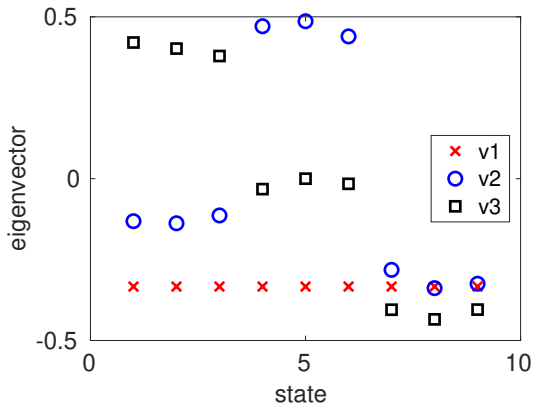
Nearly decoupled Markov chains

The structure is also present in **non-reversible** Markov chains if separating real and complex parts of the eigenvectors.

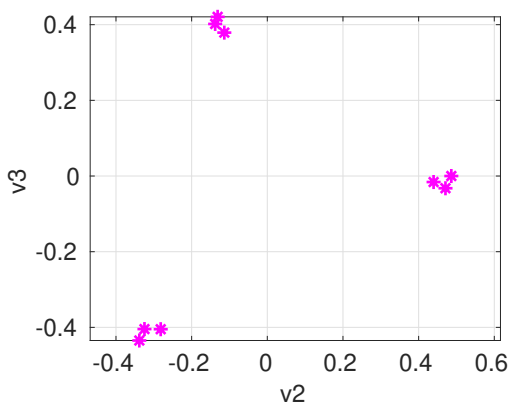
```
eps=1e-1;
P=P+eps*rand(9,9);
P=diag(1./sum(P,2))*P;  %stochasticity
```



eigenvalues



eigenvectors



simplex structure

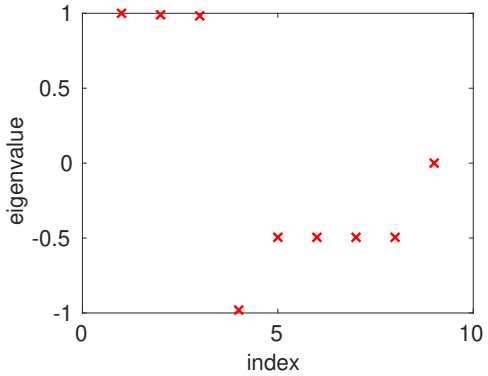
[Frank, Röblitz (2024). Spectral clustering of Markov chain transition matrices with complex eigenvalues. doi: 10.1016/j.cam.2024.115791]

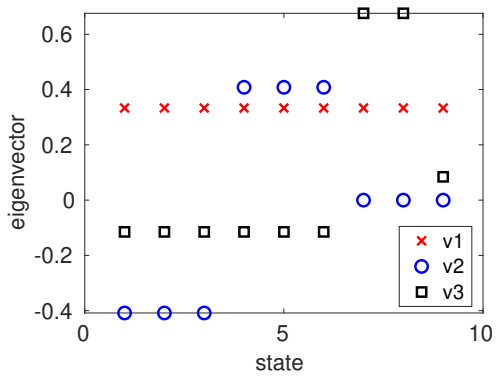
Nearly uncoupled Markov chains with transition states

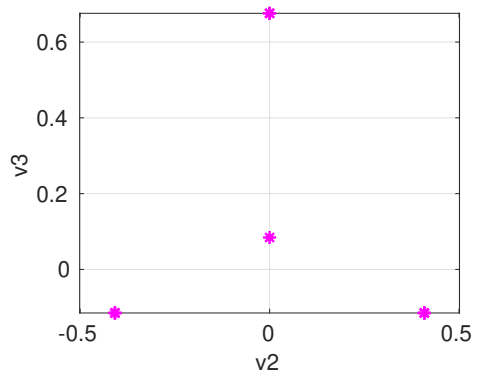
The simplex structure reveals the existence of **transition states**.

```
P(9,:)=0.1; P(:,9)=0.01;     %make state 9 a transition state
P=P+eps*rand(9,9);
P=diag(1./sum(P,2))*P
```

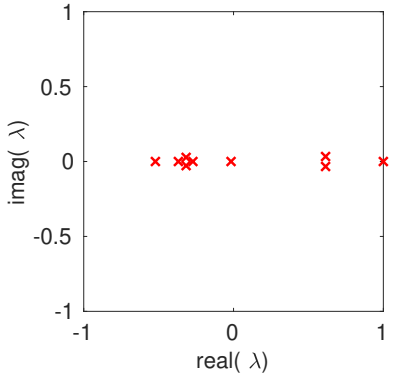
$\varepsilon = 0$:

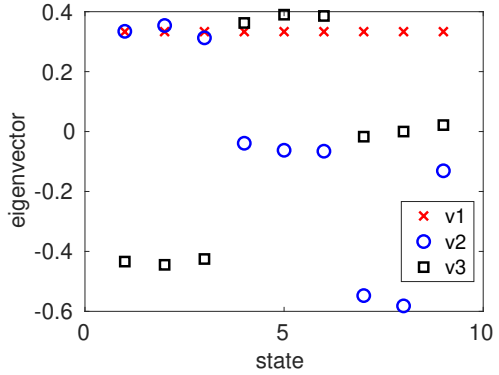


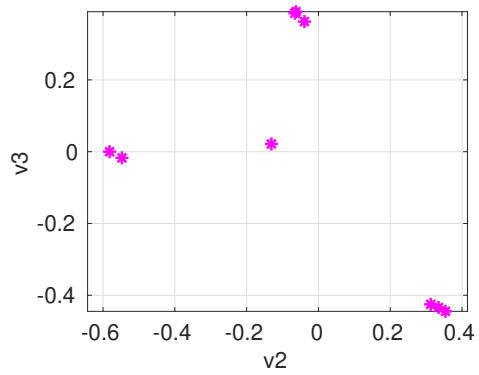




$\varepsilon = 0.1$:







eigenvalues eigenvectors simplex structure

Robust Perron Cluster Analysis (PCCA+)

Transformation of invariant subspace X to membership vectors χ:

1. $\chi = XA \quad (A \text{ regular}, PX = X\Lambda)$ (invariance)
2. $\sum_{J=1}^{n_c} \chi_J(i) = 1 \quad \forall i \in \{1, \ldots, N\}$ (partition of unity)
3. $\chi_J(i) \geq 0 \quad \forall i \in \{1, \ldots, N\},\ J \in \{1, \ldots, n_C\}$ (positivity)

In the general case (nearly uncoupled Markov chain with transition states), this problem has **no unique solution** A.

Robust Perron Cluster Analysis (PCCA+)

Transformation of invariant subspace X to membership vectors χ:

1. $\chi = XA \quad (A \text{ regular}, PX = X\Lambda)$ (invariance)
2. $\sum_{J=1}^{n_c} \chi_J(i) = 1 \quad \forall i \in \{1, \ldots, N\}$ (partition of unity)
3. $\chi_J(i) \geq 0 \quad \forall i \in \{1, \ldots, N\},\ J \in \{1, \ldots, n_C\}$ (positivity)

In the general case (nearly uncoupled Markov chain with transition states), this problem has **no unique solution** A.

Algorithmic idea of PCCA+:

- find a suitable initial guess via the inner simplex algorithm

  [M. Weber (2003). Clustering by using a simplex structure.]

- optimize A to maximize crispness:

$$n_c - \text{trace}(D_c^{-1} \chi^T D \chi) \to \min$$

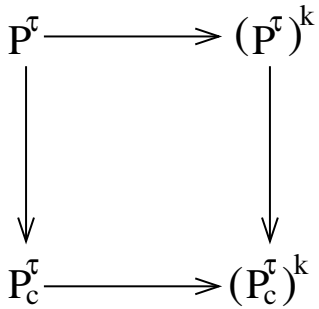
($D_c = \text{diag}(\chi^T \pi)$, $D = \text{diag}(\pi)$, $X^T D X = Id$)

[S. Röblitz, M. Weber (2013). Fuzzy spectral clustering by PCCA+. doi: 10.1007/s11634-013-0134-6]

Markov State Models

Motivation: Time-scale preserving coarse graining

$$P_c^\tau := \underbrace{\left(D_c^{-1}\chi^T D\chi\right)^{-1}}_{=:\mathcal{S}} \underbrace{\left(D_c^{-1}\chi^T DP^\tau\chi\right)}_{\text{stochastic}} = A^{-1}\Lambda_P A$$



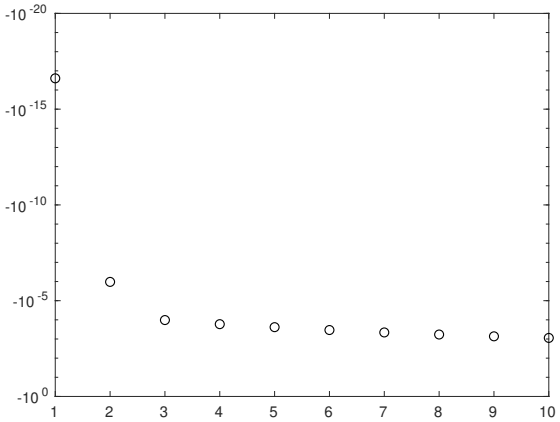
Analogously:

$$M_c := \underbrace{\left(D_c^{-1}\chi^T D\chi\right)^{-1}}_{=:\mathcal{S}} \underbrace{\left(D_c^{-1}\chi^T DM\chi\right)}_{\text{rate matrix}} = A^{-1}\Lambda_M A, \quad P_c^\tau = exp(\tau \cdot M_c)$$

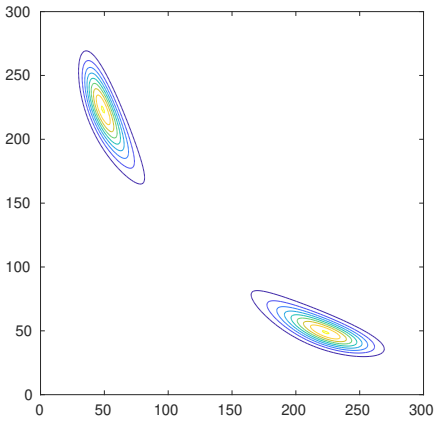
- coarse-grained rate matrix that preserves the dominant eigenvalues and hence the time scale of the slow processes
- propagation of the projected density vector commutes with the projection of the propagated density

$$\mathbf{p}^T\chi P_c^\tau = \mathbf{p}^T P^\tau \chi$$

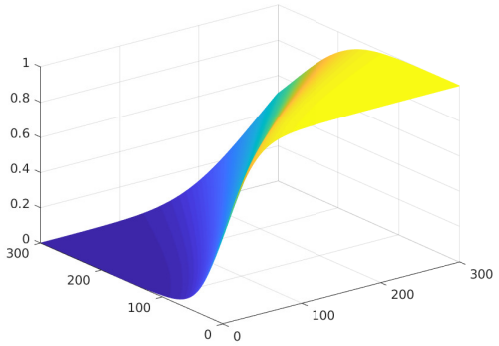
Toggle switch

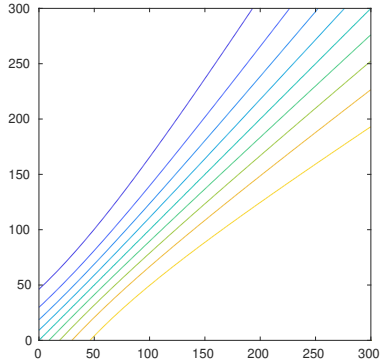


eigenvalues $\lambda_{1,\ldots,10}$

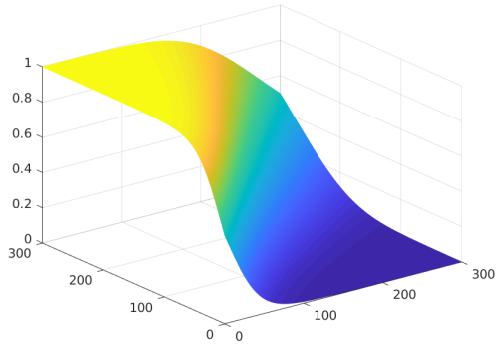


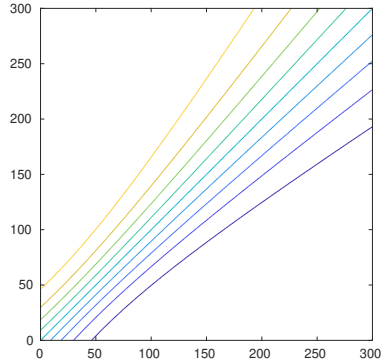
$\pi(x)$





$\chi_1(x)$



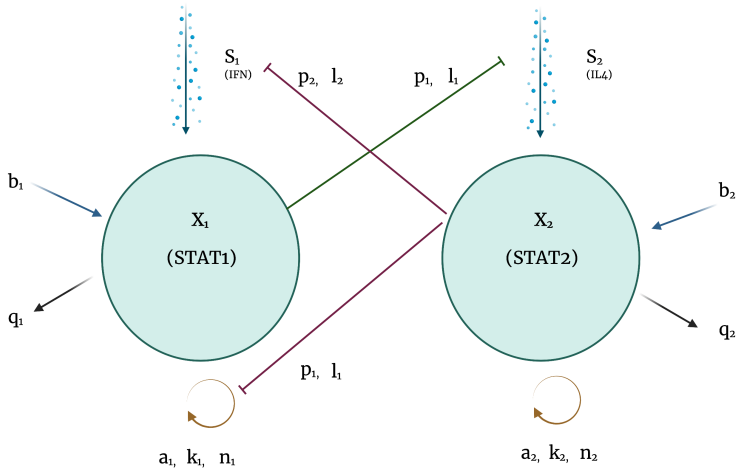


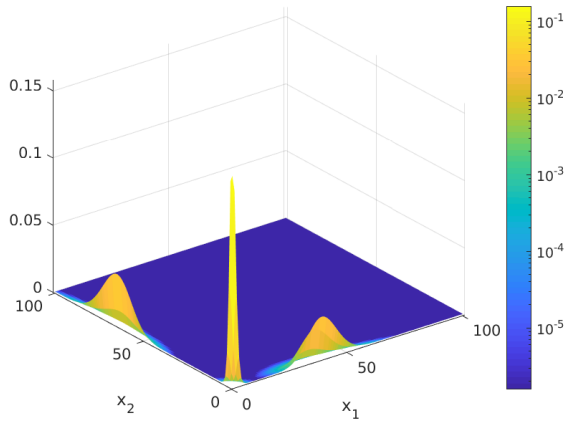
$\chi_2(x)$

statistical weights: $w = \chi^T \pi = [0.5, 0.5]^T$

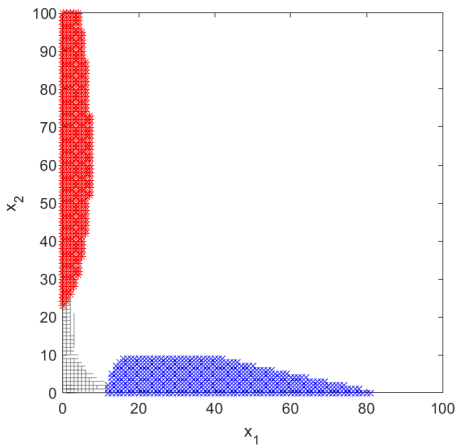
$$P_c(\tau = 5000) = \begin{pmatrix} 0.9974 & 0.0026 \\ 0.0026 & 0.9974 \end{pmatrix}$$

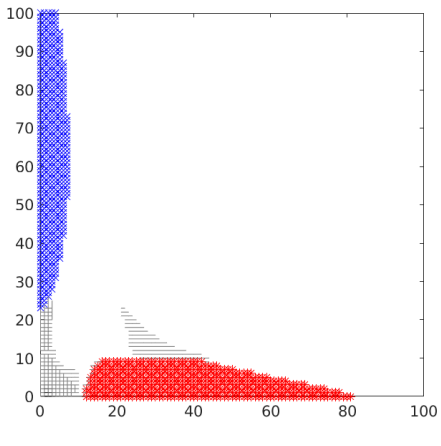
Phenotype transitions in macrophage polarization





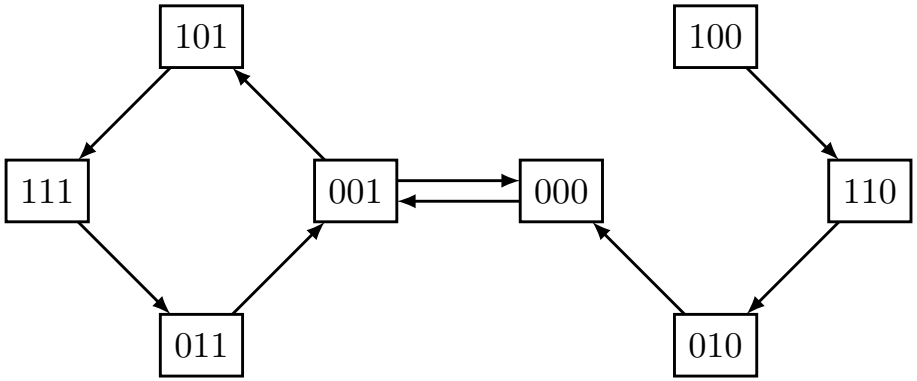
Discrete transition path theory (TPT) for Markov jump processes [Metzner et al. 2009]:

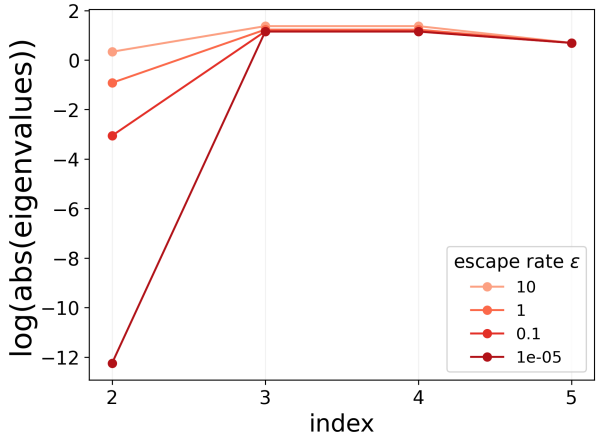




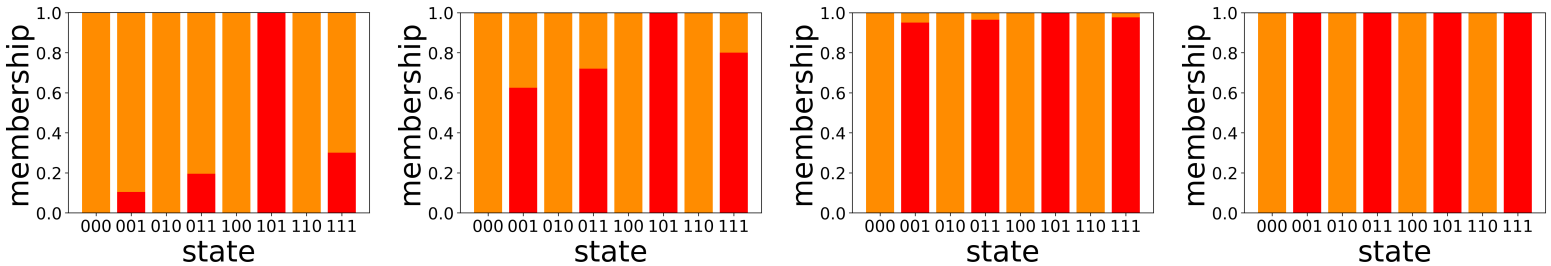
[Frank et al. (2022). Macrophage phenotype transitions in a stochastic gene-regulatory network model. DOI: 10.1016/j.jtbi.2023.111634]

Continuous-time Boolean models



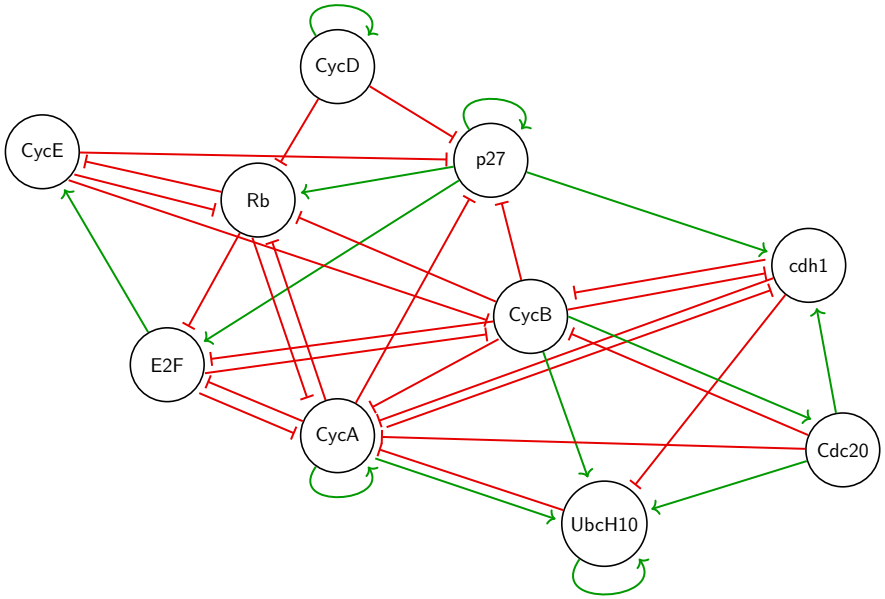


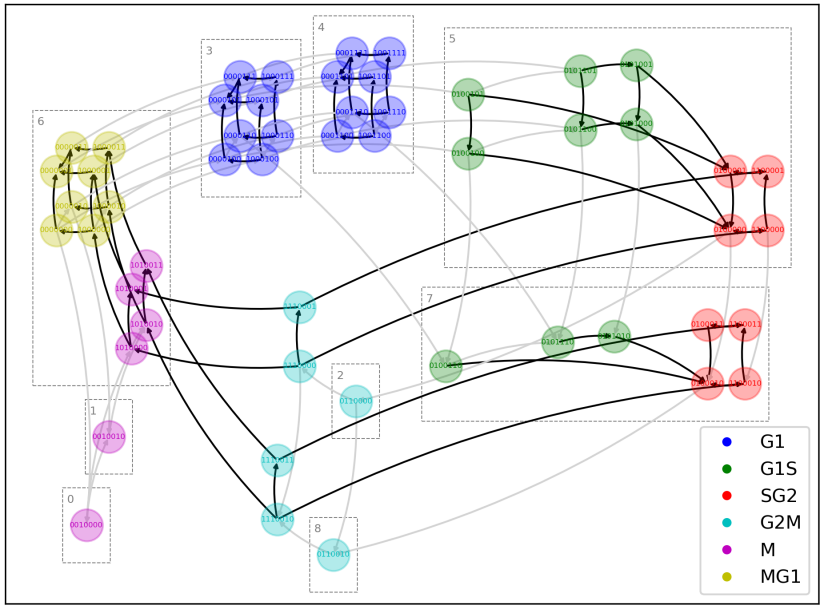
$$K = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \varepsilon & -(Au+\varepsilon) & 0 & 0 & 0 & Au & 0 & 0 \\ Bd & 0 & -Bd & 0 & 0 & 0 & 0 & 0 \\ 0 & Bd & 0 & -Bd & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -Bu & 0 & Bu & 0 \\ 0 & 0 & 0 & 0 & 0 & -Bu & 0 & Bu \\ 0 & 0 & Ad & 0 & 0 & 0 & -Ad & 0 \\ 0 & 0 & 0 & Ad & 0 & 0 & 0 & -Ad \end{pmatrix}.$$

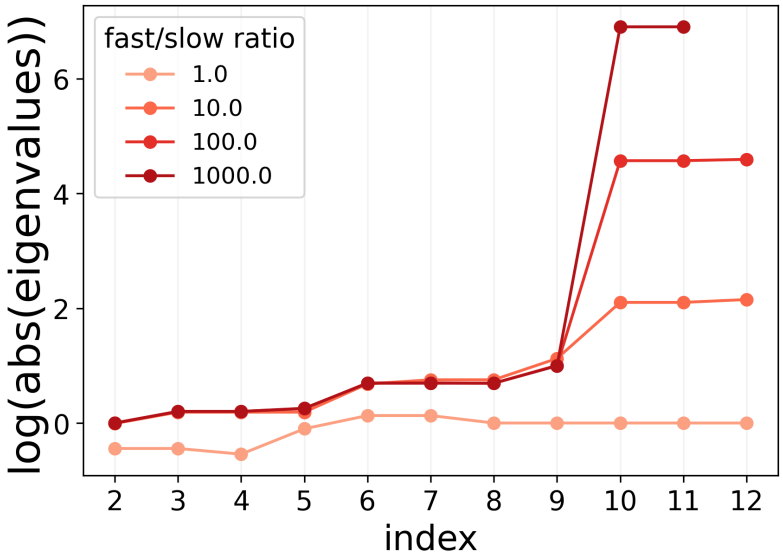


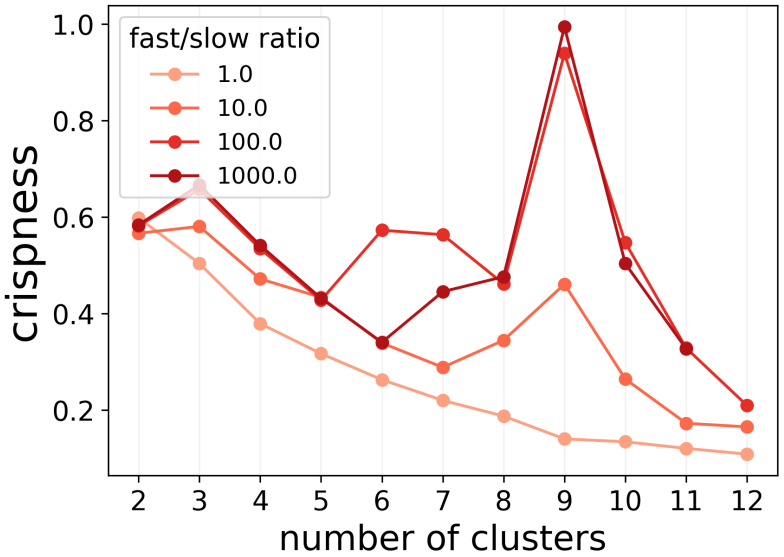
[Yousefian et al. (2024). DOI: 10.1007/978-3-031-71671-3_16]

Mammalian cell cycle model (Fauré et al. 2006)





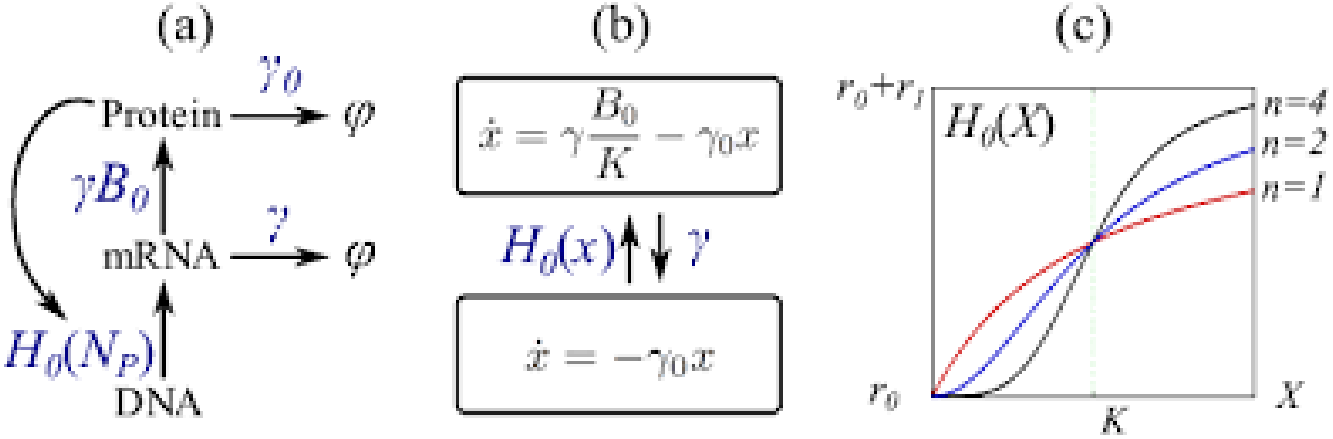




[Yousefian et al. (2024). DOI: 10.1007/978-3-031-71671-3_16]

Piecewise-deterministic Markov processes (PDMP)

Coarse-grained model describing only the protein population while fully accounting for the effects of discrete and fluctuating mRNA population:

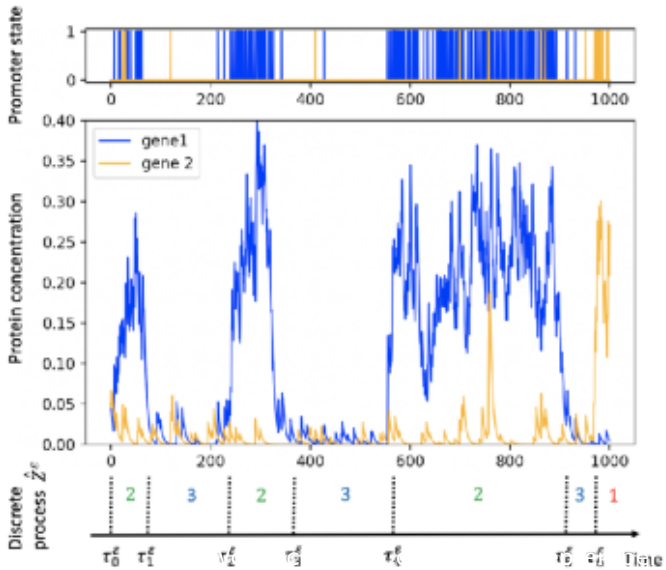


[Lin et al. 2016, DOI: 10.1103/PhysRevE.93.022409; Ventre et al. 2021, DOI: 10.1007/s00285-021-01684-1]

Master equation:

$$\frac{\partial}{\partial t}\begin{bmatrix} p_1(x,t) \\ p_0(x,t) \end{bmatrix} = L^\dagger \begin{bmatrix} p_1(x,t) \\ p_0(x,t) \end{bmatrix}$$

$$L^\dagger = \begin{bmatrix} -\gamma - \partial_x(\gamma b - \gamma_0 x) & H(x) \\ \gamma & -H(x) + \gamma_0 \partial_x x \end{bmatrix}$$



Drawback

- the size of M increases exponentially with the number of species
- even if M is sparse, bookkeeping of its entries requires enumeration of all reachable states

Drawback

- the size of M increases exponentially with the number of species
- even if M is sparse, bookkeeping of its entries requires enumeration of all reachable states

We need some discretization strategy!

Discretization

$$\mathcal{M}^* C_k(x) = \lambda_k C_k(x), \quad \lambda_k \approx 0$$

Given a partition of unity $\{\psi_i(x)\}_{i=1}^N$, the metastable functions $C_k(x)$ can be approximated by

$$C_k(x) \approx C_k^N(x) = \sum_{i=1}^{N} \chi_{ik} \psi_i(x).$$

The condition $\sum_k \chi_{ik} = 1$ is sufficient to ensure that $\{C_k^N(x)\}$ is a partition of unity if $\{\psi_i(x)\}_{i=1}^N$ is one.

$\Rightarrow$ the vectors $\chi_k = (\chi_{1k}, \ldots, \chi_{Nk})^T$ can be interpreted as **membership vectors**

Inner products with test functions $\{\phi_j(x)\}_{i=1}^N \rightarrow$ discrete eigenvalue problem

$$Q\chi_k = \lambda_k S \chi_k$$

with

$$Q_{ij} \equiv \langle \phi_i, \mathcal{M}^* \psi_j \rangle = \langle \mathcal{M} \phi_i, \psi_j \rangle \quad \text{and} \quad S_{ij} \equiv \langle \phi_i(x), \psi_j(x) \rangle.$$

Algorithmic approach

Choose $\{\psi_i(x)\}_{i=1}^N$ to be a Voronoi tesselation (meshless!) of $\mathbb{N}^D$:

$$Q_{ij} \equiv \langle \phi_i, \mathcal{M}^* \psi_j \rangle = \sum_{x \in \mathbb{N}^D} \sum_{r=1}^{R} \alpha_r(x)[\psi_j(x+\nu_r) - \psi_j(x)]\phi_i(x)$$

$\Rightarrow$ requires the detection of all states x at the boundary of ψ_j from where the support of ψ_j can be left within one reaction step

$\Rightarrow$ very difficult (impossible) on unstructured Voronoi tesselations in high dimensions

Alternatives:

- use of overlapping (e.g. radial) basis function: How to combine SSA with importance sampling?
- switch to transition probabilities

From transition rates to transition probabilities

By the Hille-Yosida theorem, $\mathcal{M}$ is the infinitesimal generator of a strongly continuous semigroup $\mathcal{T}^{(\tau)}$ and the CME admits a unique continuous solution in the form

$$p(x, t+\tau) = \mathcal{T}^{(\tau)} p(x, t).$$

Galerkin discretization of $\mathcal{T}^{(\tau)}$ in terms of Voronoi cell $\{\psi_i(x)\}_{i=1}^N$ and test functions $\phi_i(x) = \psi_i(x)\pi(x) \equiv \pi_i(x)$:

$$P\chi_k = \lambda_k \chi_k$$

with

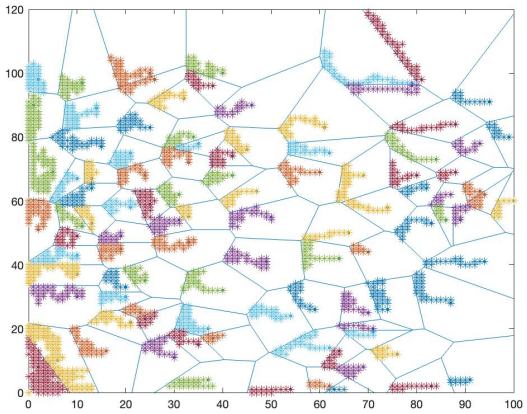
$$P_{ij}^{(\tau)} \equiv \frac{\langle \phi_i, (\mathcal{T}^{(\tau)})^* \phi_j \rangle_\pi}{\langle \phi_i, \phi_i \rangle_\pi} = \frac{\sum_{x\in\Omega_i} \sum_{y\in\Omega_j} \pi(x) T^{(\tau)}(x,y)}{w_i}, \quad i,j = 1,\ldots,N$$

If we have sampled points $\{x_k\}_{k=1}^K$ according to the partial stationary density $\pi_i(x)$, then P_{ij} can be approximated by

$$P_{ij} \approx \frac{1}{K} \sum_{k=1}^{K} \sum_{y\in\Omega_j} T^{(\tau)}(x_k, y).$$

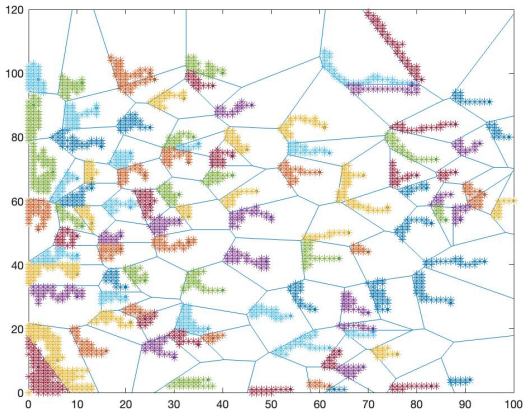
Markov State Model building

1. Horizontal sampling

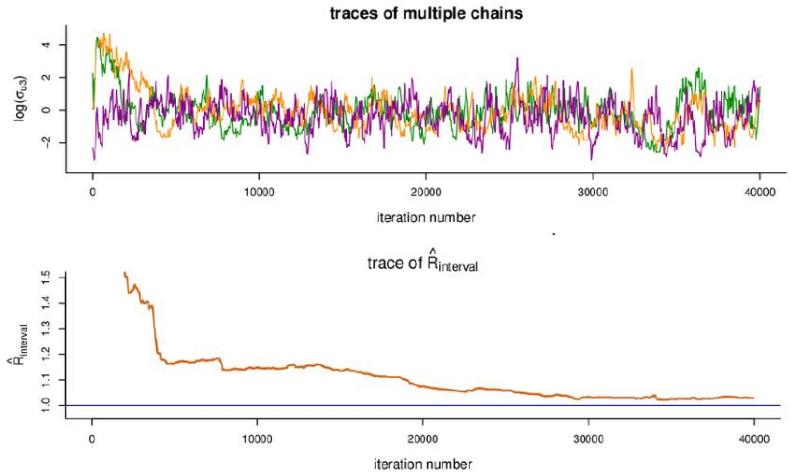


Markov State Model building

1. Horizontal sampling

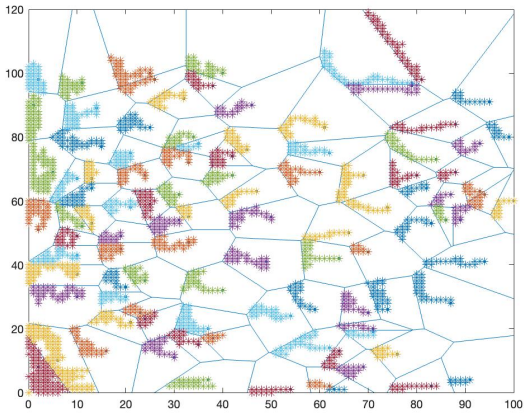


2. Convergence analysis

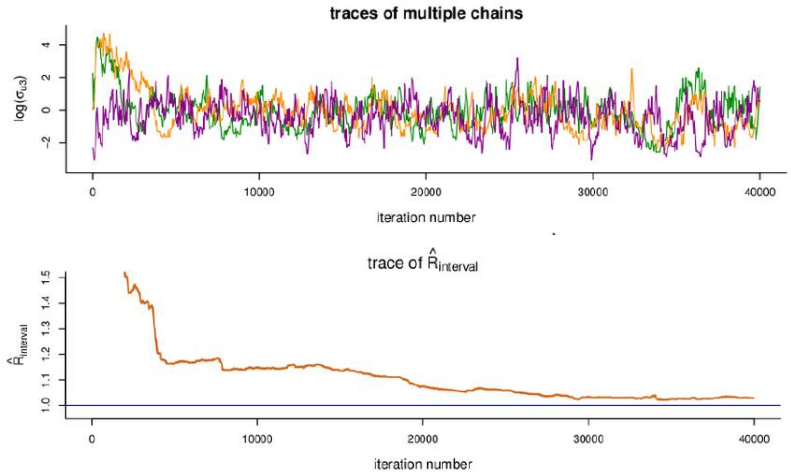


Markov State Model building

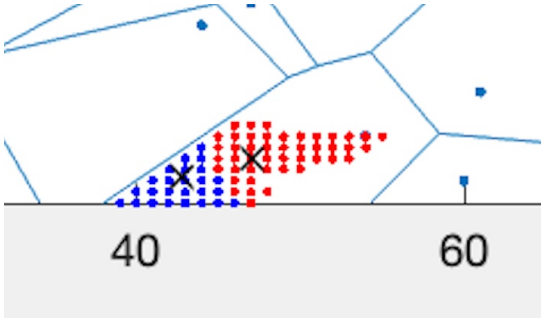
1. Horizontal sampling



2. Convergence analysis

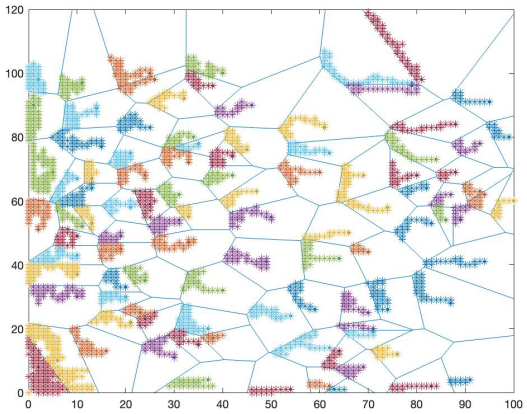


3. Hierarchical refinement

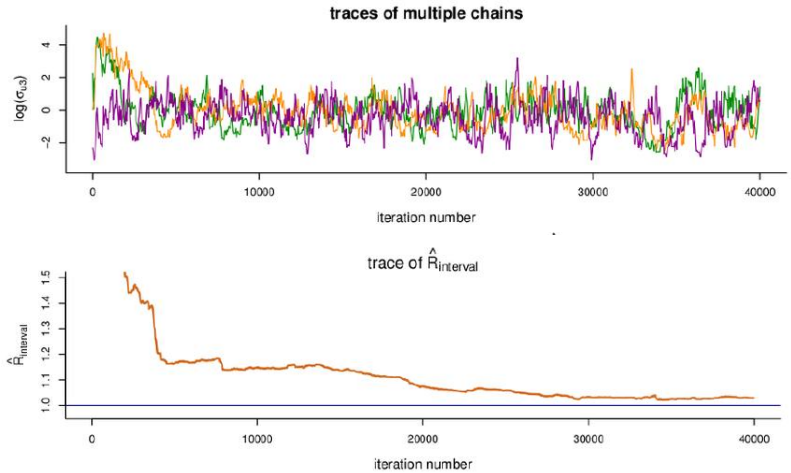


Markov State Model building

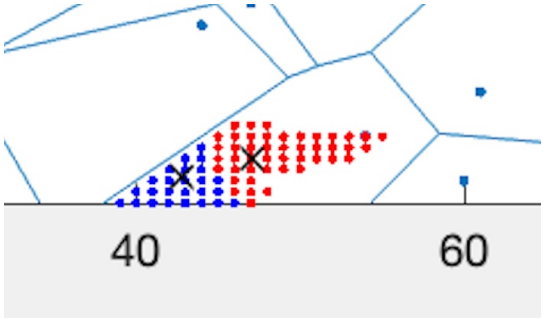
1. Horizontal sampling



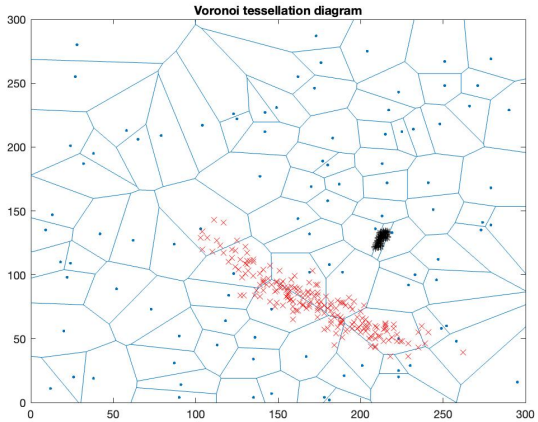
2. Convergence analysis



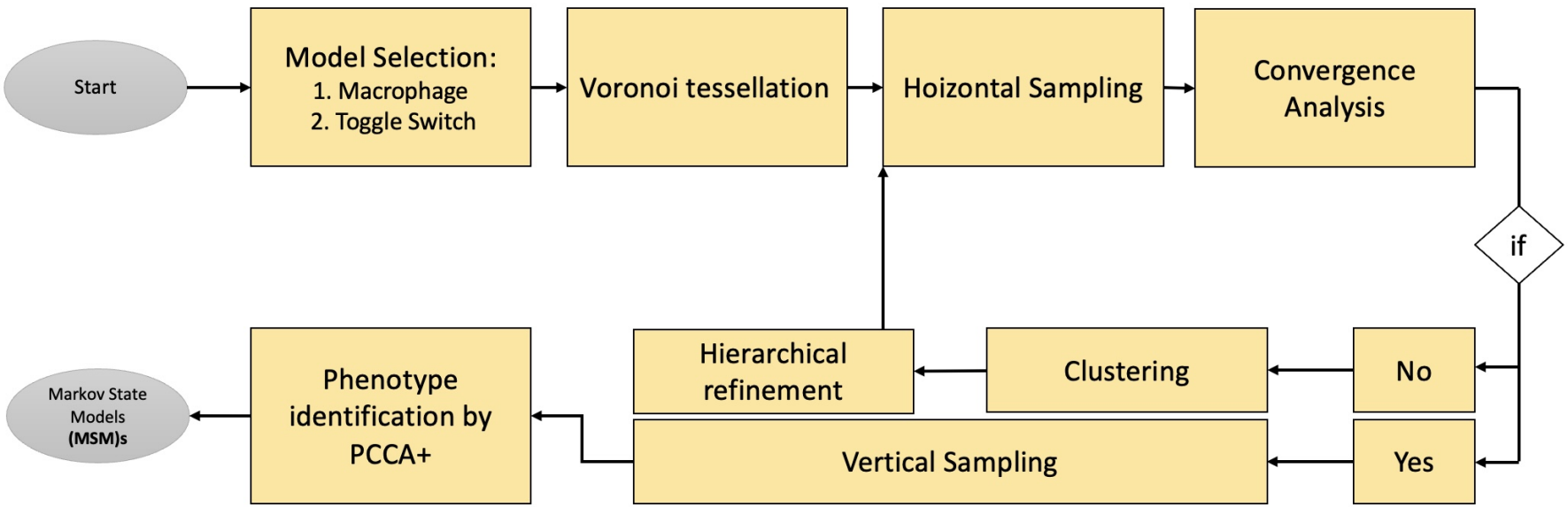
3. Hierarchical refinement



4. Vertical sampling

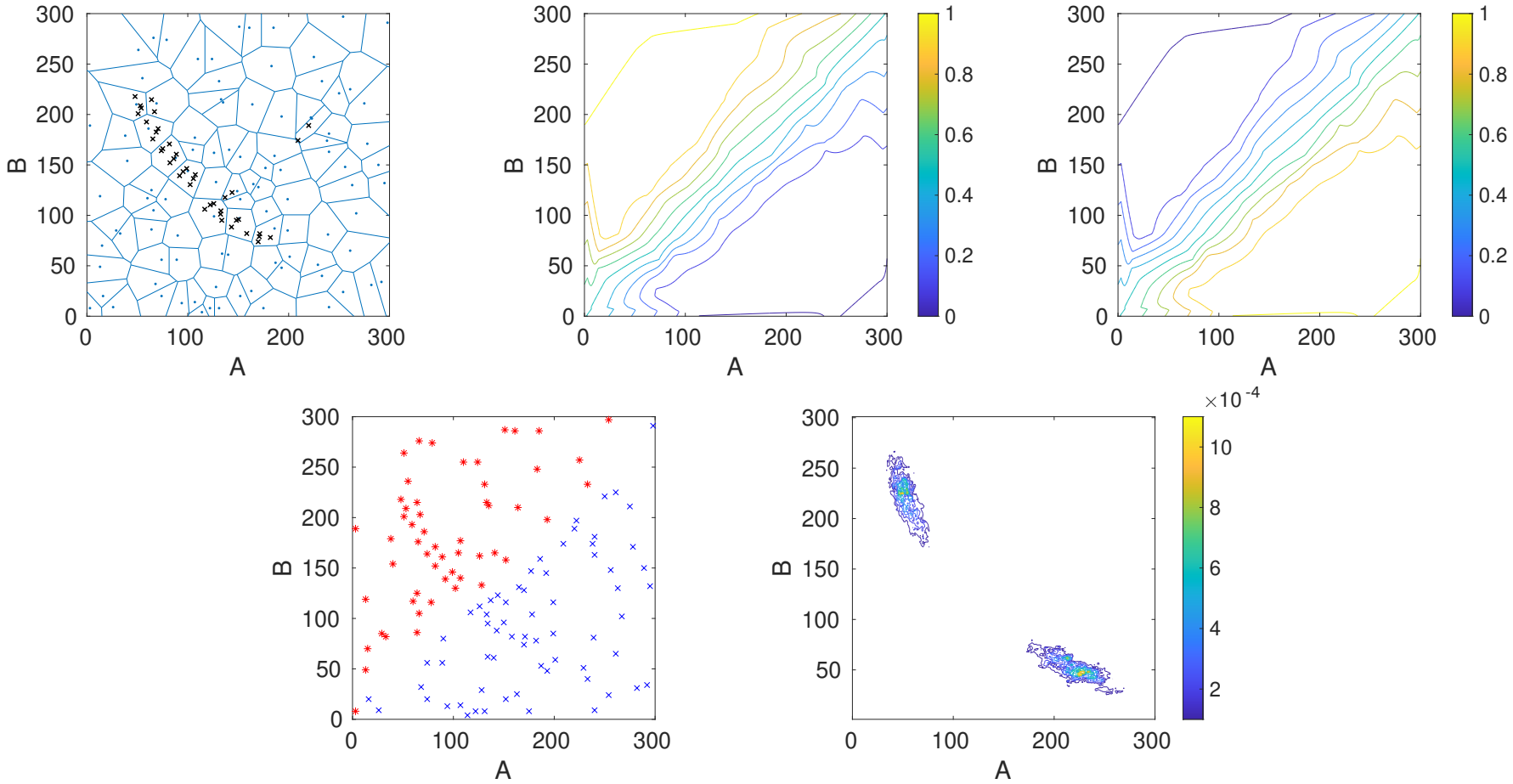


Markov State Model building

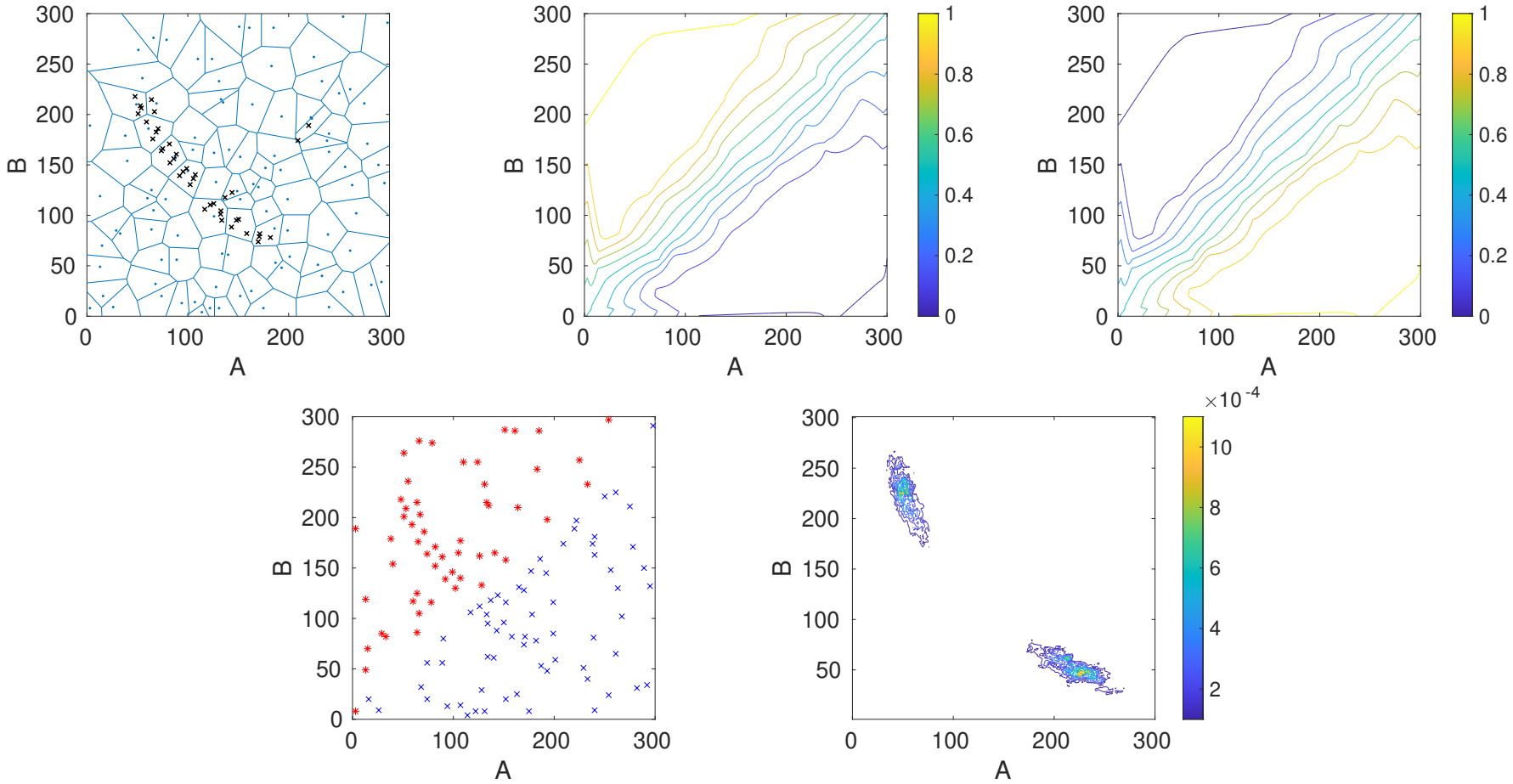


[M. Yousefian, A.-S. Frank, M. Weber, S. Röblitz (2024). Efficient construction of Markov state models for stochastic gene regulatory networks by domain decomposition. *bioarXiv*. doi: 10.1101/2023.11.21.568127]

Toggle Switch: Results



Toggle Switch: Results



Error estimation/Uncertainty quantification?

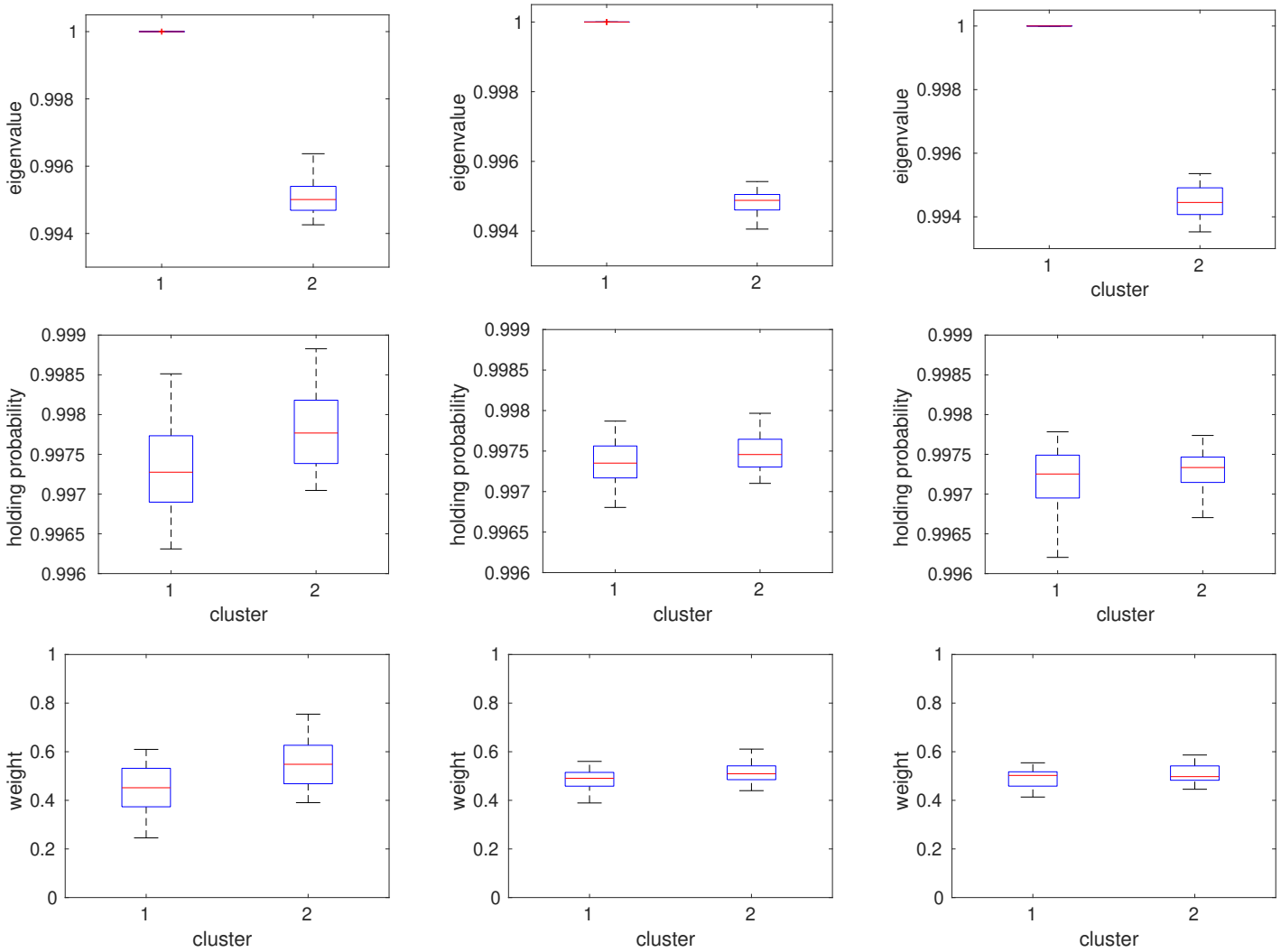
Uncertainty quantification

The rows of the transition probability matrix P are statistically independent.

- for each cell, repeat the vertical sampling multiple times by randomly picking a fixed number of points from the horizontal sampling
- this results in multiple candidates for each row $P(i,:)$
- the distribution of probability vectors $P(i,:)$ follows the **Dirichlet distribution** $\mathrm{Dir}(\alpha)$
- estimate α by maximum-likelihood estimation
- sample transition probability matrices P and compute multiple MSMs to quantify the uncertainty

Toggle switch: Uncertainty quantification

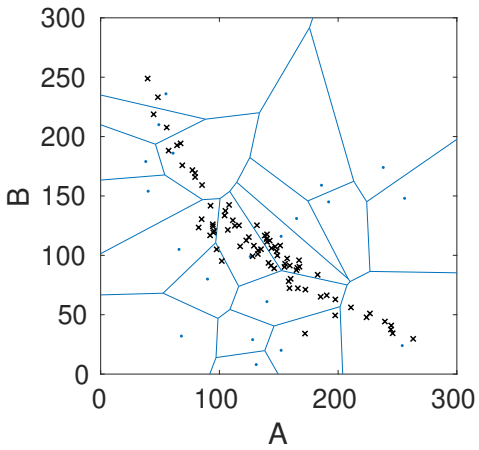
Increasing the number of vertical sampling points reduces the uncertainty:

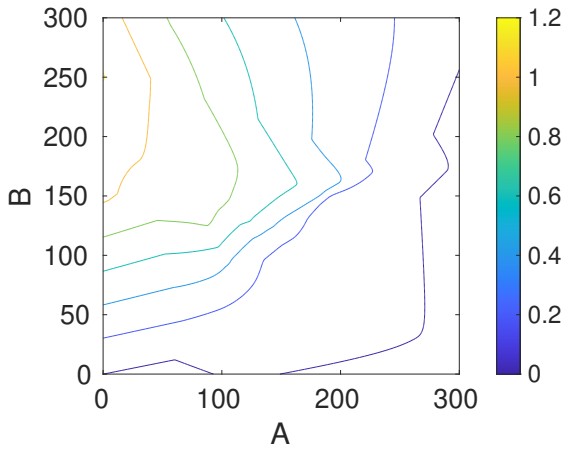


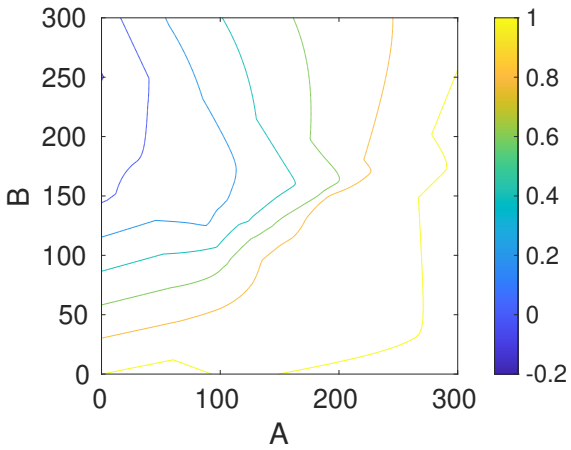
$N_0 = 100,\ N_{\text{vert}} = 100$ $N_0 = 100,\ N_{\text{vert}} = 500$ $N_0 = 20,\ N_{\text{vert}} = 500$

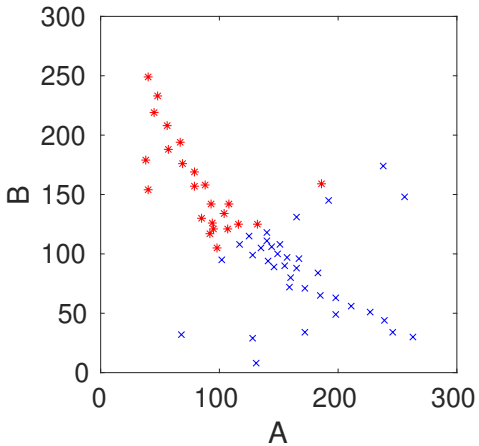
Toggle Switch: Results

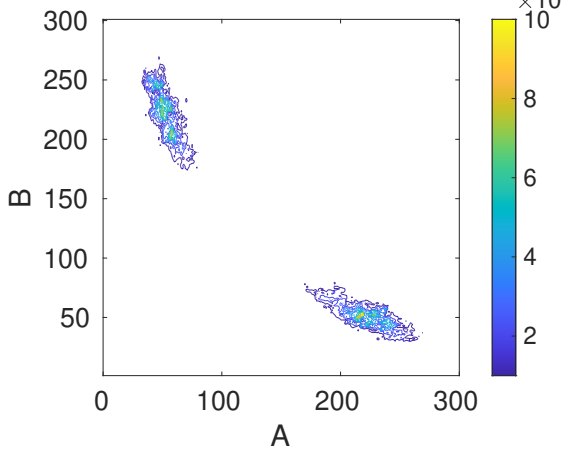
Decreasing the number of Voronoi cells does not deteriorate the discretization error:











Summary

Conclusion

- fluctuations in gene expression do not just average away but can instead lead to easily detectable differences between otherwise identical cells
- Markov state modelling, developed for molecular conformation dynamics, can be **generalized** to other multi-stable stochastic dynamical systems, including **non-reversible GRNs**
- for small model systems with finite state space, PCCA+ can directly be applied to the transfer operator or its generator
- more complex model systems require **discretization**

Summary

Conclusion

- fluctuations in gene expression do not just average away but can instead lead to easily detectable differences between otherwise identical cells
- Markov state modelling, developed for molecular conformation dynamics, can be **generalized** to other multi-stable stochastic dynamical systems, including **non-reversible GRNs**
- for small model systems with finite state space, PCCA+ can directly be applied to the transfer operator or its generator
- more complex model systems require **discretization**

Future work

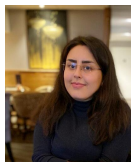
- test other discretization methods for the CME (e.g. radial basis function)
  $\rightarrow$ importance sampling if stationary distribution is unknow?
- discretization strategies for operators other than the CME
  $\rightarrow$ how to decompose the Boolean state space?
- software development (parallelization $\rightarrow$ higher dimensions, GUI, SBML)
- avoid discretization by "learning" the eigenfunctions using an artificial neural network (ISOKANN)

  [R. J. Rabben et al. (2020). ISOKANN: Invariant subspaces of Koopman operators learned by a neural network. doi: 10.1063/5.0015132]

Key publications

- M. Yousefian, E. Tonello, A.-S. Frank, H. Siebert, S. Röblitz. Uncovering dynamic structures within cyclic attractors of asynchronous Boolean networks with spectral clustering. In: Computational Methods in Systems Biology. CMSB 2024. *Lecture Notes in Computer Science*, vol 14971. https://doi.org/10.1007/978-3-031-71671-3_16
- M. Yousefian, A.-S. Frank, M. Weber, S. Röblitz. Efficient construction of Markov state models for stochastic gene regulatory networks by domain decomposition. Preprint available on bioRxiv. https://doi.org/10.1101/2023.11.21.568127
- A.-S. Frank, K. Larripa, H. Ryu, S. Röblitz. Macrophage phenotype transitions in a stochastic gene-regulatory network model. *Journal of Theoretical Biology* 575:111634, 2023. https://doi.org/10.1016/j.physbeh.2022.114034
- A.-S. Frank, K. Larripa, H. Ryu, R. G. Snodgrass, S. Röblitz. Bifurcation and sensitivity analysis reveal key drivers of multistability in a model of macrophage polarization. *Journal of Theoretical Biology*, 2020. https://doi.org/10.1016/j.jtbi.2020.110511
- S. Röblitz and M. Weber. Fuzzy spectral clustering by PCCA+: Application to Markov state models and data classification. *Advances in Data Analysis and Classification* 7(2):147–179, 2013. https://doi.org/10.1007/s11634-013-0134-6

Thank you for your attention!



Maryam Yousefian

Anna-Simone Frank

With funding from

The Research Council of Norway

Collaborators:

Marcus Weber (ZIB Berlin)

Heike Siebert (U Kiel)

Kamilla Larripa (Humboldt State U, CA, US)

Hwayeon Ryu (Elon University, NC, US)

TROND MOHN FOUNDATION

American Institute of Mathematics