Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

D-BSSE
Department of Biosystems
Science and Engineering

Probabilistic Methods for Biochemical Reaction Networks

Mustafa Khammash
Department of Biosystems Science and Engineering
ETH Zürich

LECTURE III

Outline

- Computing the probability density function
- Computing the statistical moments
- Moment-based inference
- Density-based inference
- Future directions

Computing the Probability Density Function

Solving the Master Equation

Computing the PDF

Goal: Compute $p(x,t)$, the probability that $X(t) = x$.

The Chemical Master Equation (CME):

$$\frac{dp(x,t)}{dt} = -p(x,t)\sum_k w_k(x) + \sum_k p(x - s_k, t) w_k(x - s_k)$$

Enumerate the state space: $\mathcal{X} = \{x_1, x_2, x_3, \ldots\}$

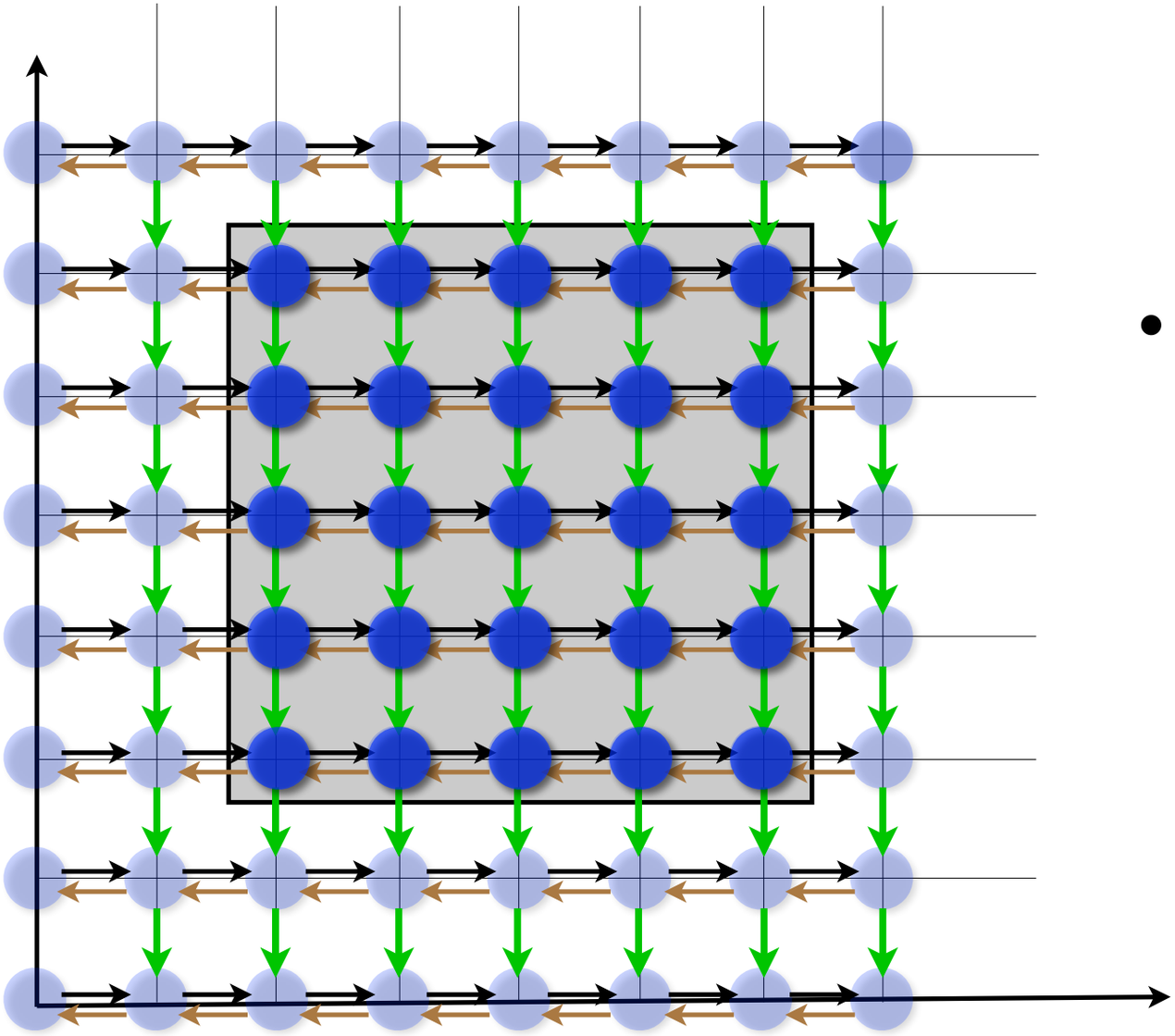
Form the probability density state vector $P(\mathcal{X}, \cdot) : R^+ \to \ell_1$

$$P(\mathcal{X}, t) := [p(x_1, t) \quad p(x_2, t) \quad p(x_3, t) \quad \ldots \quad]^T$$

CME can now be written in matrix form:

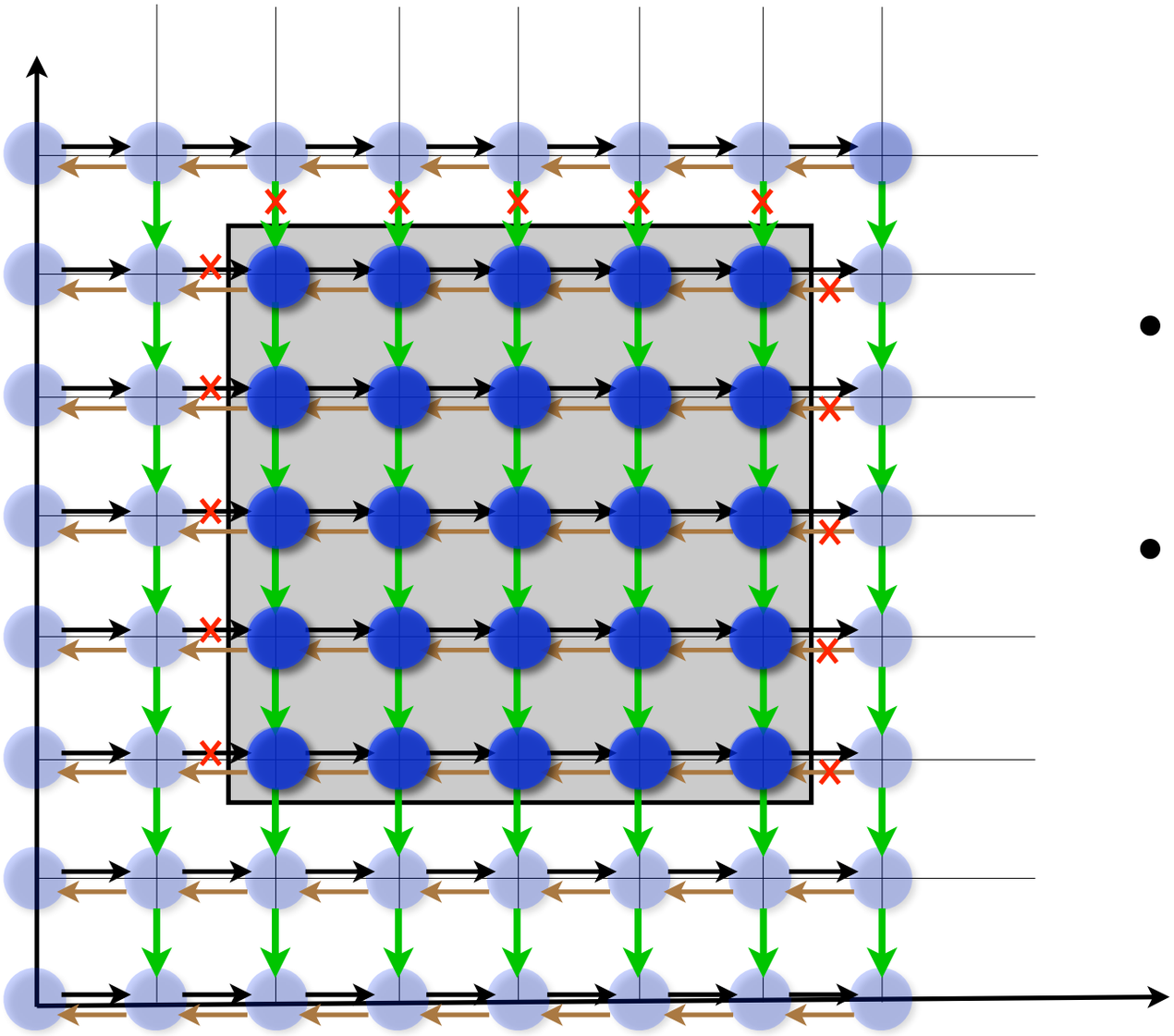
$$\dot{P}(\mathcal{X}, t) = \mathbf{A} \cdot P(\mathcal{X}, t)$$

The Finite State Projection Approach



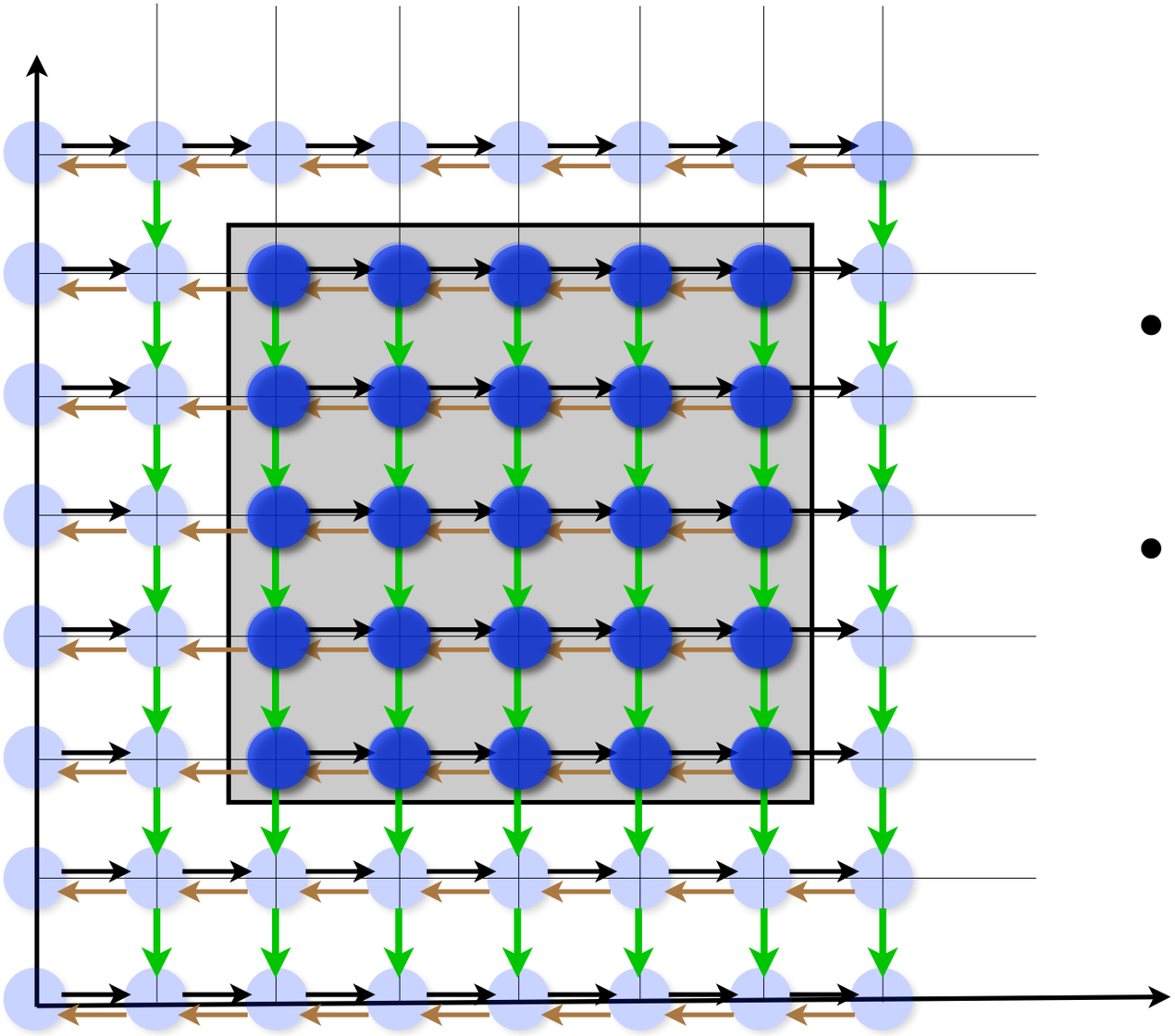
- A finite subset is appropriately chosen

The Finite State Projection Approach



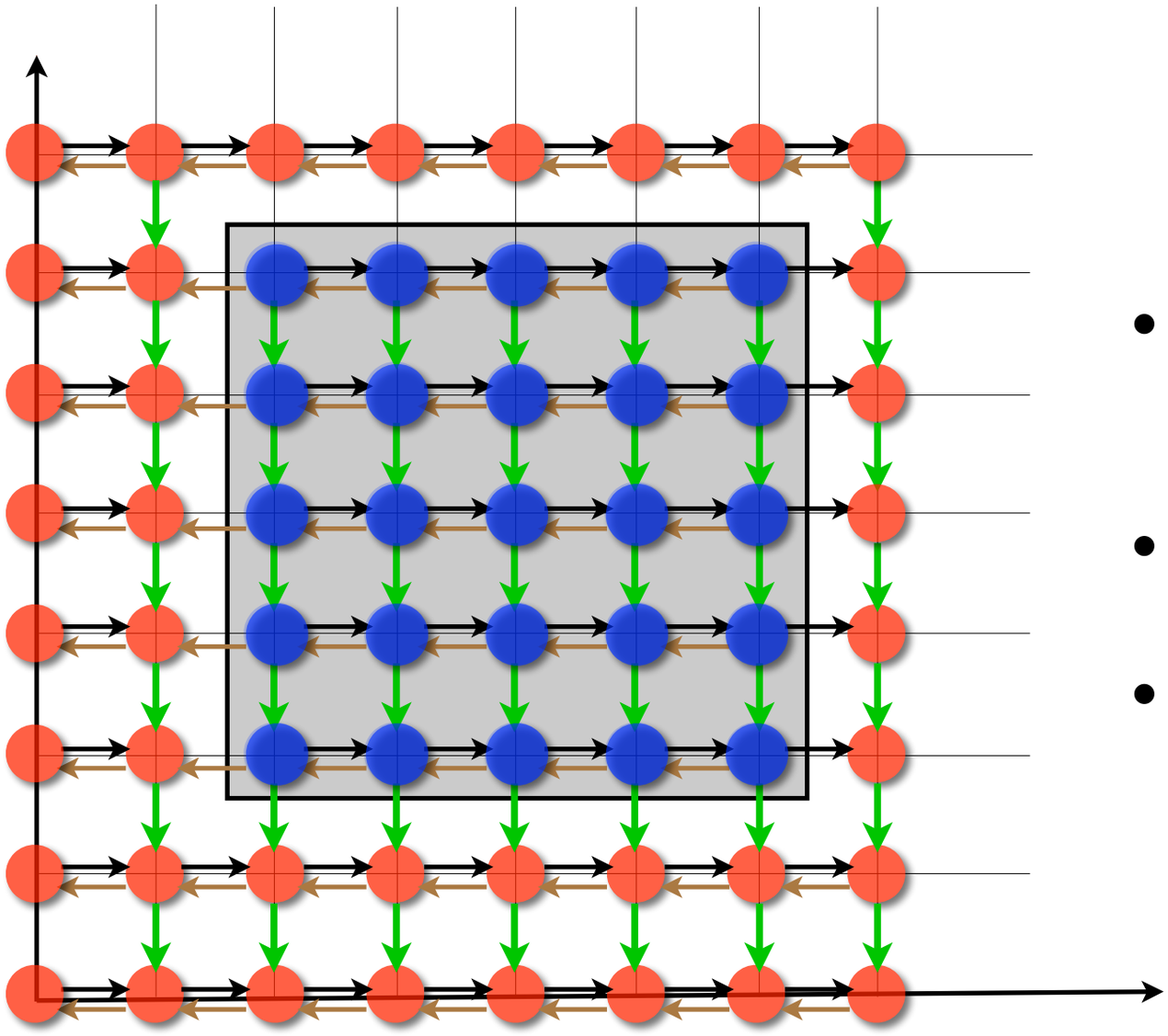
- A finite subset is appropriately chosen
- Transitions into subset are deleted

The Finite State Projection Approach



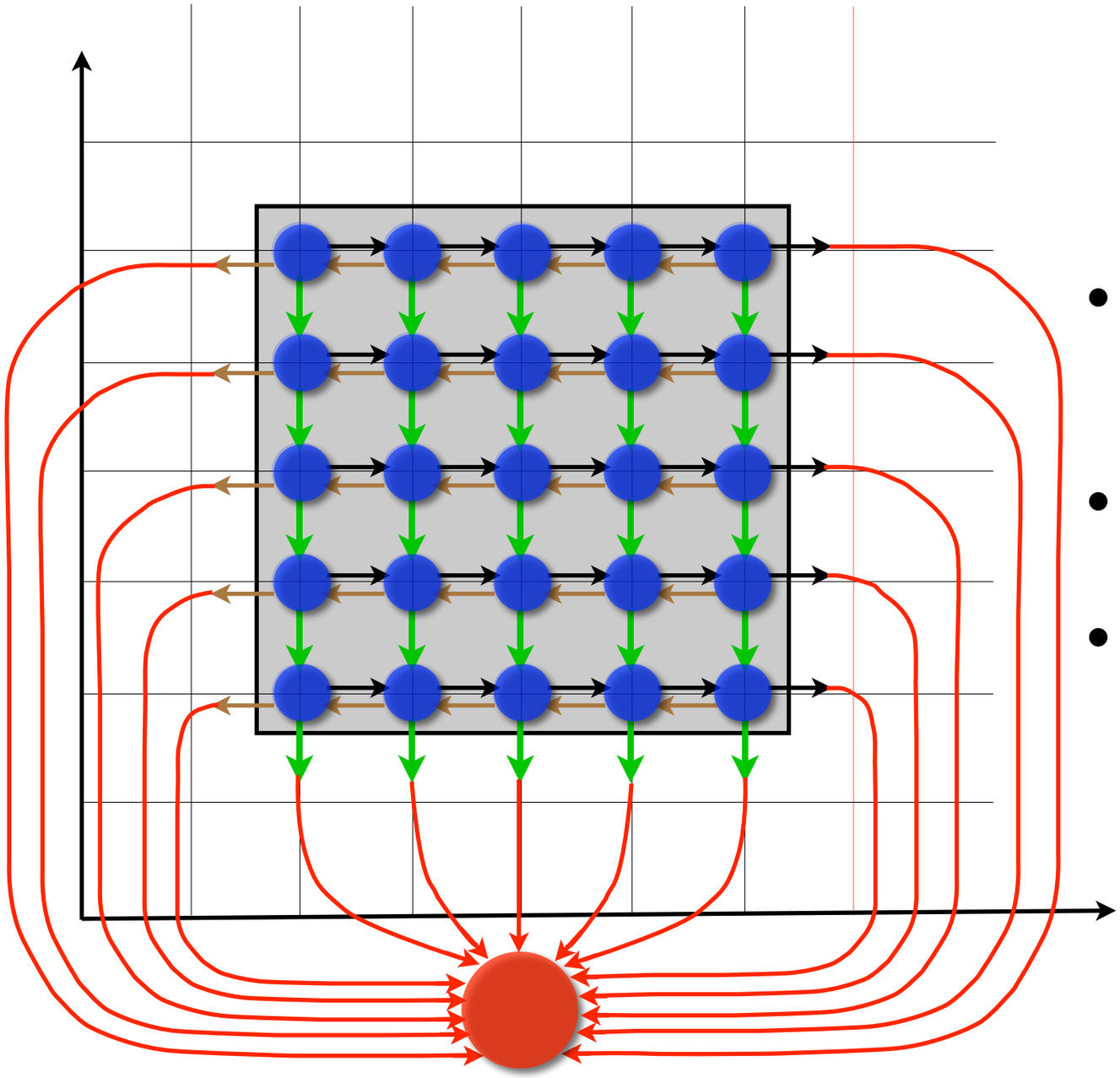
- A finite subset is appropriately chosen
- Transitions into subset are deleted

The Finite State Projection Approach



- A finite subset is appropriately chosen
- Transitions into subset are deleted
- Remaining (infinite) states are projected into a single state

The Finite State Projection Approach



- A finite subset is appropriately chosen
- Transitions into subset are deleted
- Remaining (infinite) states are projected into a single state

The projected system can be solved exactly!

Guaranteed Error Bounds

Notation: For a matrix A, let A_J to be the principle submatrix of A indexed by J, where $J = [m_1 \dots m_n]$.

Projection Error Bounds Consider any Markov process described by the Forward Kolmogorov Equation:

$$\dot{P}(\mathcal{X}; t) = A \cdot P(\mathcal{X}; t).$$

If for an indexing vector J: $\mathbf{1}^T \exp(A_J T) P(\mathcal{X}_J; 0) \geq 1 - \epsilon$, then

$$\left\| \begin{bmatrix} P(\mathcal{X}_J; t) \\ P(\mathcal{X}_{J'}; t) \end{bmatrix} - \begin{bmatrix} \exp(A_J t) P(\mathcal{X}_J; 0) \\ 0 \end{bmatrix} \right\|_1 < \epsilon \qquad t \in [0, T]$$

The FSP Algorithm

---

given initial prob. vector $P(\mathcal{X}, 0)$; final time T; error tol. $\epsilon > 0$

select an initial finite subset $\mathcal{X}_J$ that includes support of $P(\mathcal{X}, 0)$

repeat

1. *expand projection*: $\mathcal{X}_J := \mathcal{X}_J \cup \mathcal{X}_{J'}$
2. *form projected system*: A_J
3. *compute probability for projection*: $\exp(A_J T) P(\mathcal{X}_J; 0)$

until $1^T \exp(A_J T) P(\mathcal{X}_J; 0) \geq 1 - \epsilon$

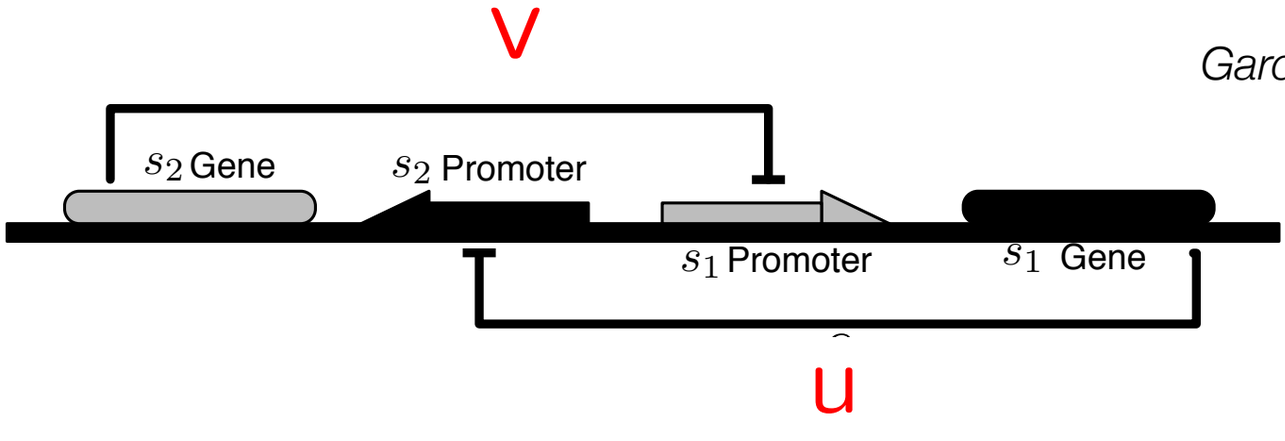
end

Certificate: $\|P(\mathcal{X}_J, T) - \exp(A_J T) P(\mathcal{X}_J; 0)\|_1 \leq \epsilon$

Application to a Stochastic Switch

Gardner, et al., Nature, 2000

Two repressors
u and **v**



u and **v** inhibit each other

$$\phi \xrightarrow{w_1} \mathbf{u} \qquad \phi \xrightarrow{w_2} \mathbf{v}$$

$$w_1 = \frac{\alpha_1}{1 + v^\beta} \qquad w_2 = \frac{\alpha_2}{1 + u^\gamma}$$

$$s_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad s_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

u and **v** degrade

$$\mathbf{u} \xrightarrow{w_3} \phi \qquad \mathbf{v} \xrightarrow{w_4} \phi$$

$$w_3 = u \qquad w_4 = v$$

$$s_3 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \qquad s_4 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

A Sample Trajectory

Trajectory simulation using Gilespie's SSA

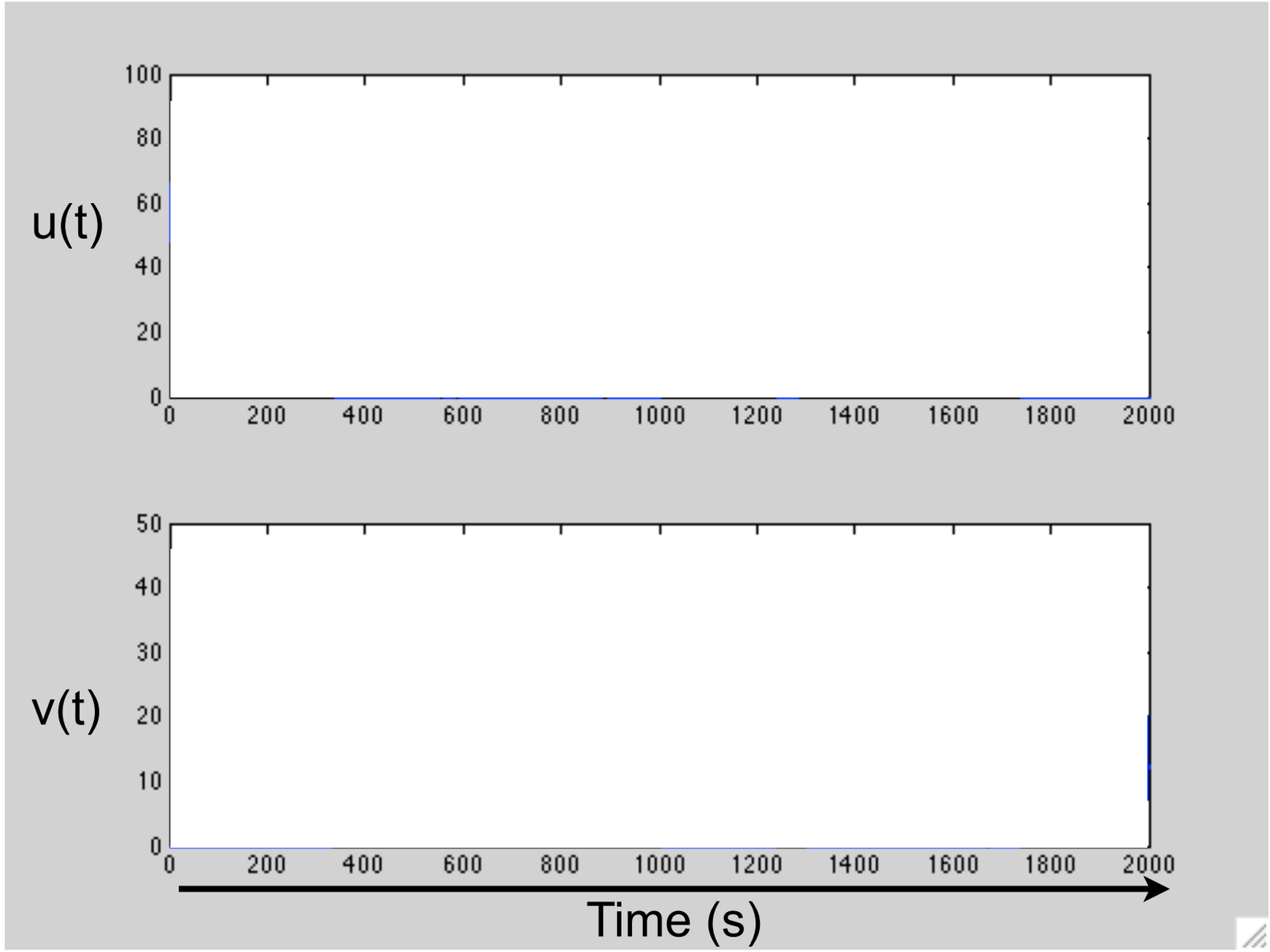
Initial conditions

$$\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \begin{bmatrix} 60 \\ 0 \end{bmatrix}$$

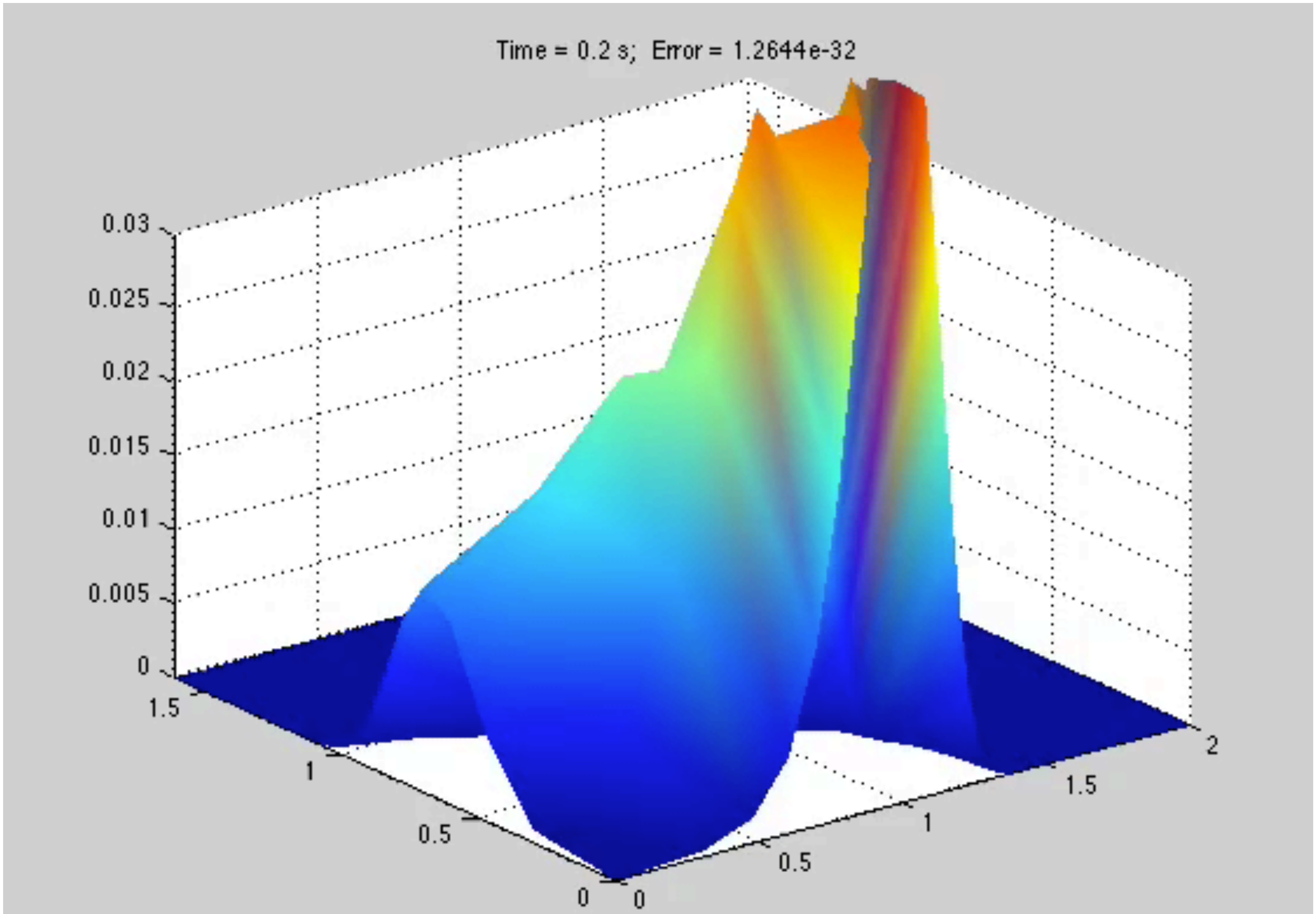
Parameters

$\alpha_1 = 50 \qquad \beta = 2.5$

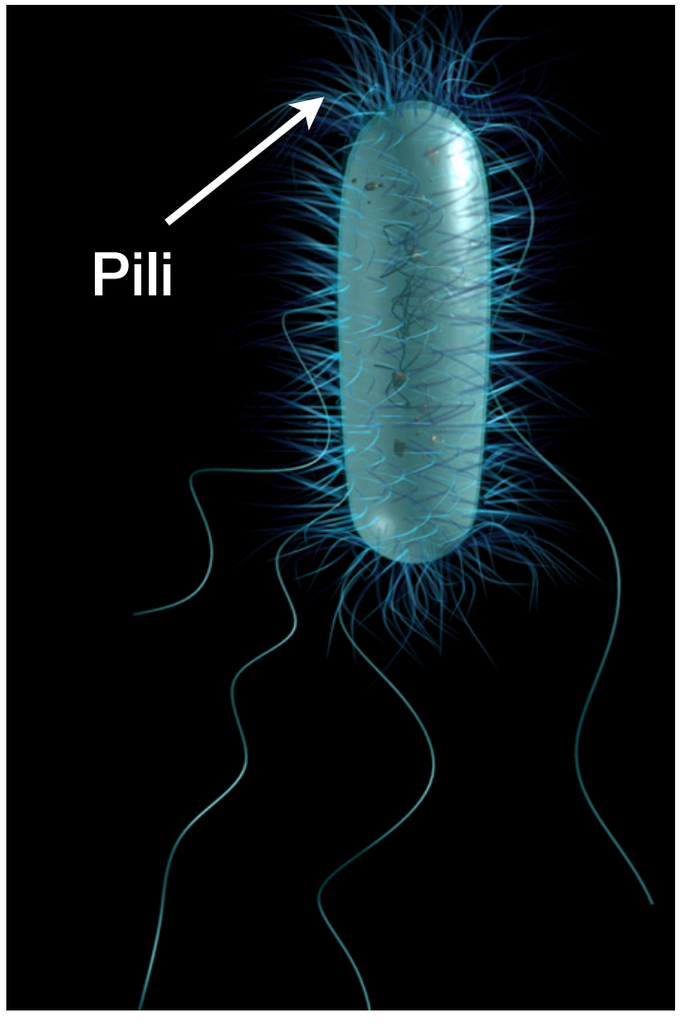
$\alpha_2 = 16 \qquad \gamma = 1$

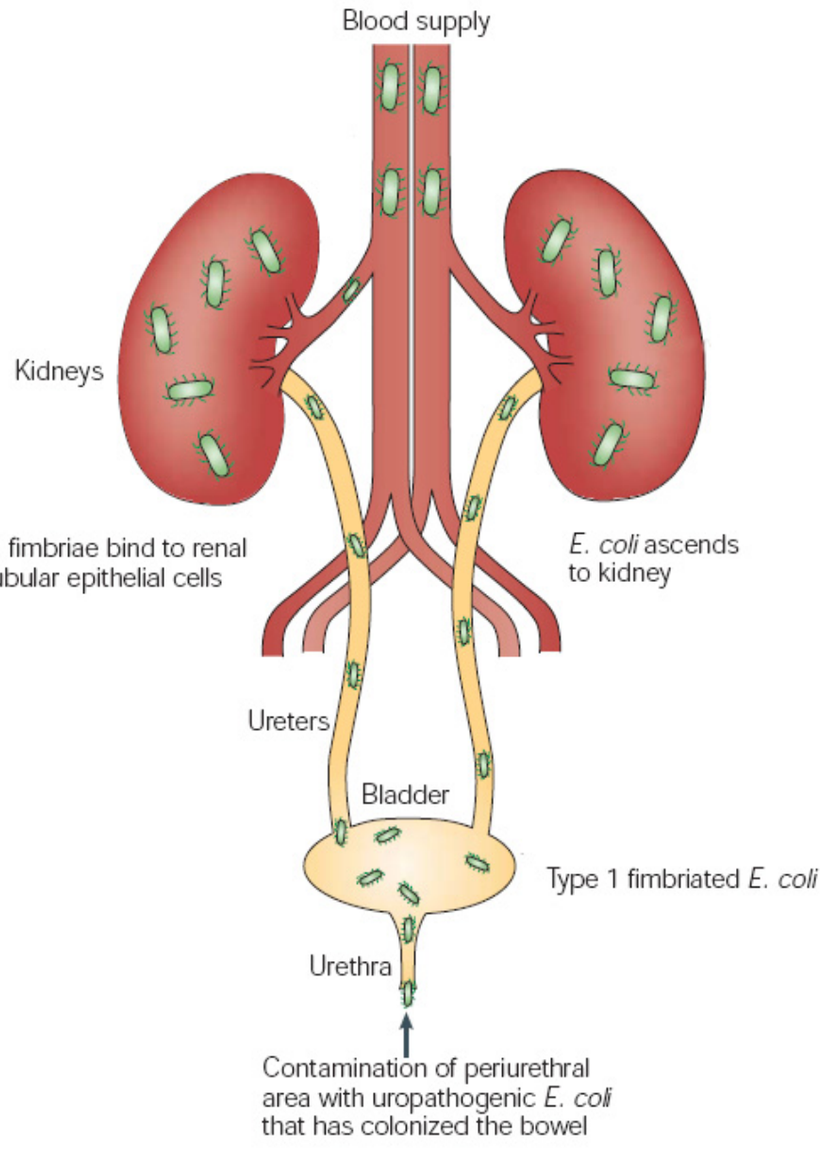


Joint pdf Computation using FSP Algorithm

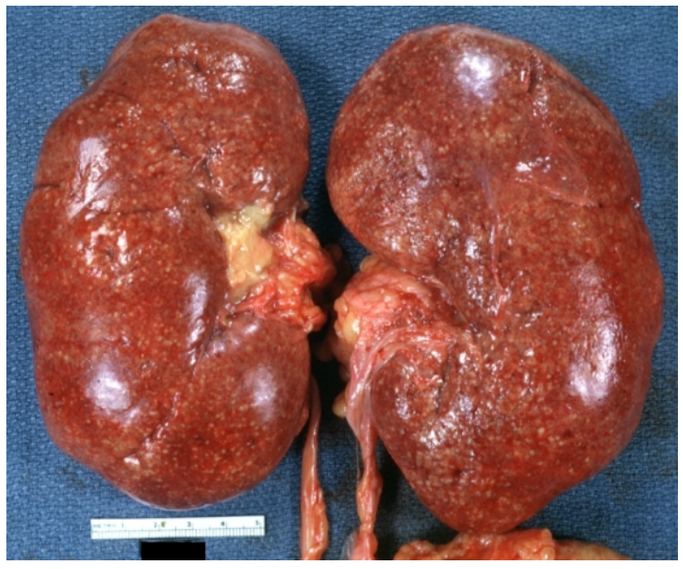


A Stochastic Switch Involved in Disease: Pyelonephritis Assoicated Pili (PAP)

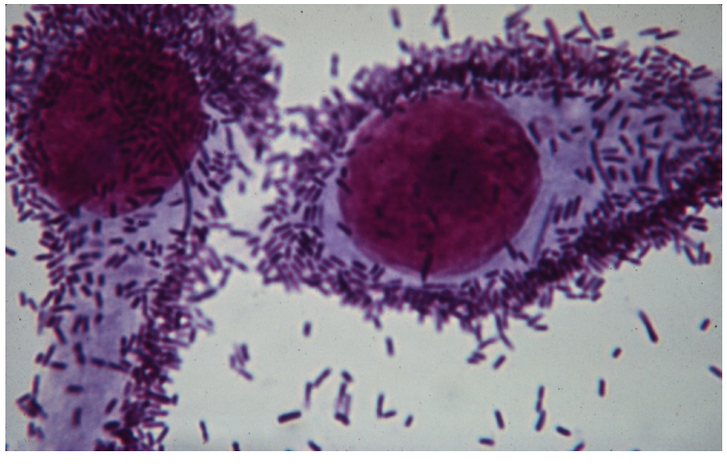




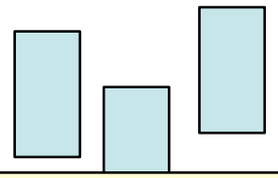
Credit: *J.B. Kaper et al Nature Rev. Microbiol (2004) 2, 123-140 (modified)*



Credit: *University of Alabama at Birmingham, Department of Pathology*



Application of FSP to the Pap Switch Model



Lrp

What is the probability of being in State g_2 at time T?

State g_2

ON

State g_3

OFF

R_5 R_6 R_7 R_8

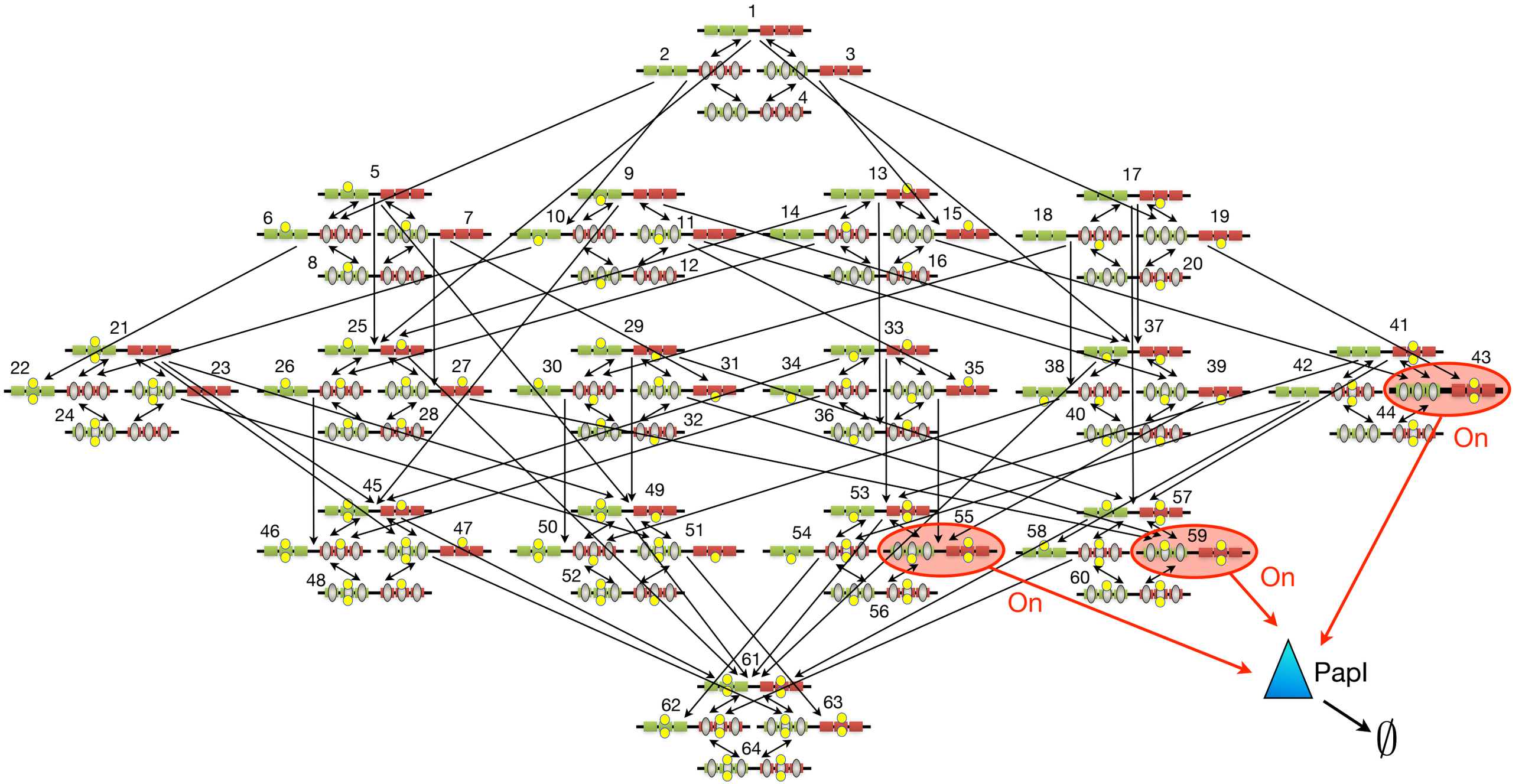
Piliation takes place if gene is ON at a specific time: T

State g_4 OFF

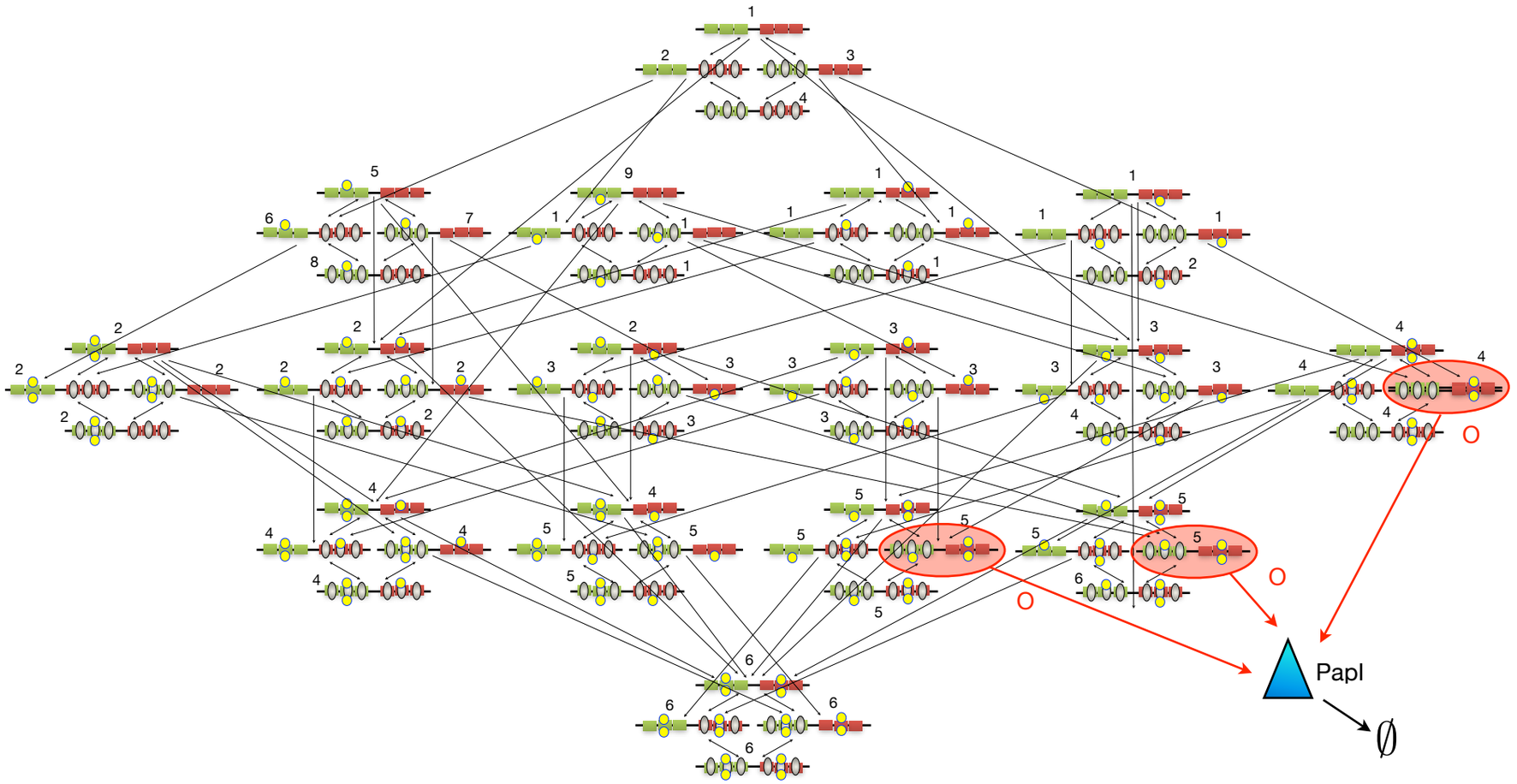
Accuracy and Efficiency

Method	Number of simulations	Time (s)	Relative error in switch rate (%)
FSP	Does not apply[a]	<4	<0.5
SSA[b]	1.25×10^5	≈18	38.8
SSA	2.5×10^5	≈35	27.3
SSA	5.0×10^5	≈70	9.9
SSA	10.0×10^5	≈140	8.5

A More Detailed Stochastic model of the pap Switch

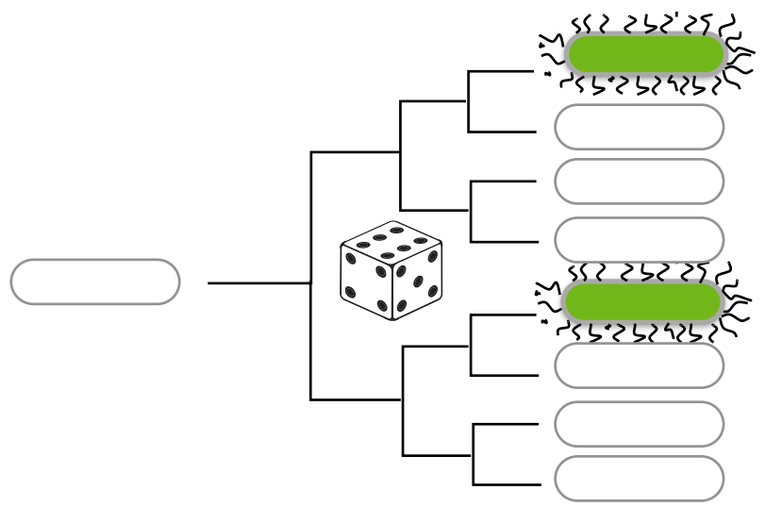


Solving The Master Equation Using FSP



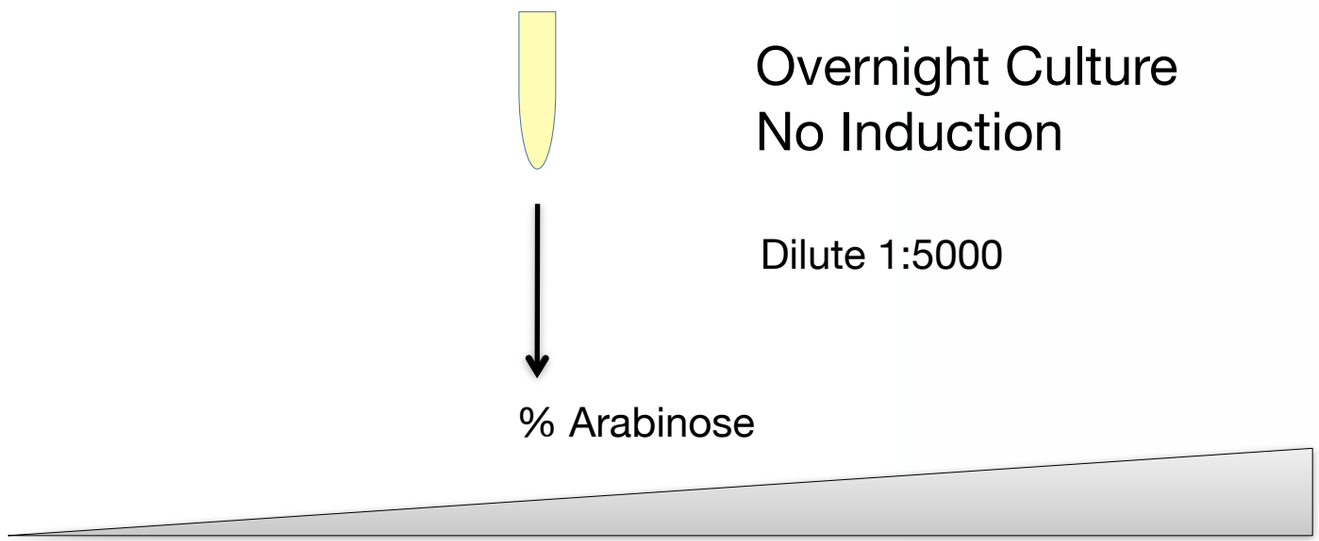
Starting from any initial state at time t=0, the probability of finding a cell in a given state(s) at t=T can be accurately computed:

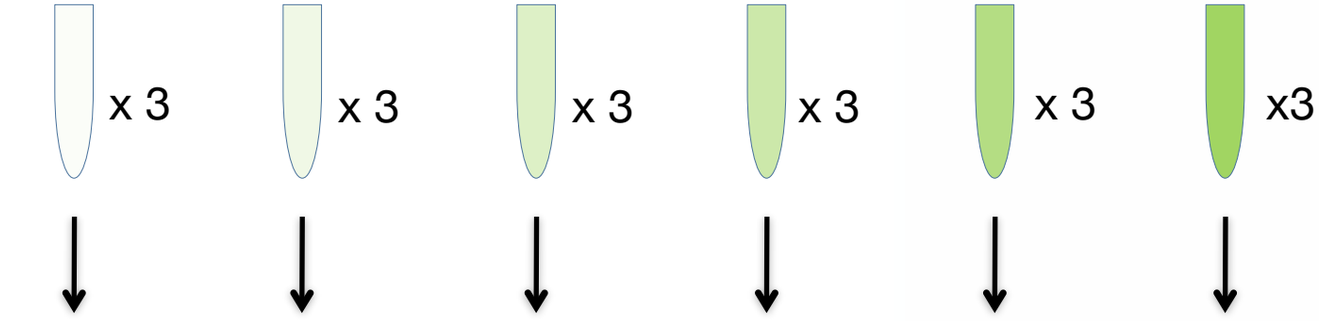
$$p(\mathcal{X}_J, T) = \exp(A_J T) p(\mathcal{X}_J, 0)$$

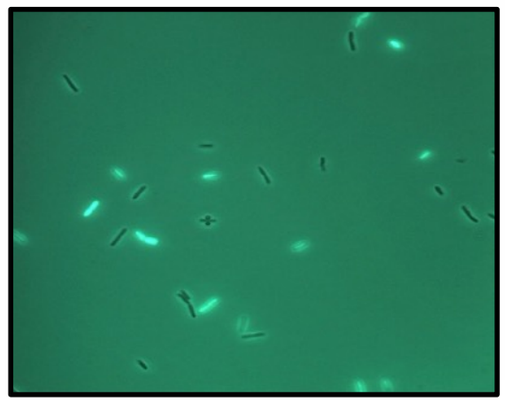


Allows us to compute probability of an ON state

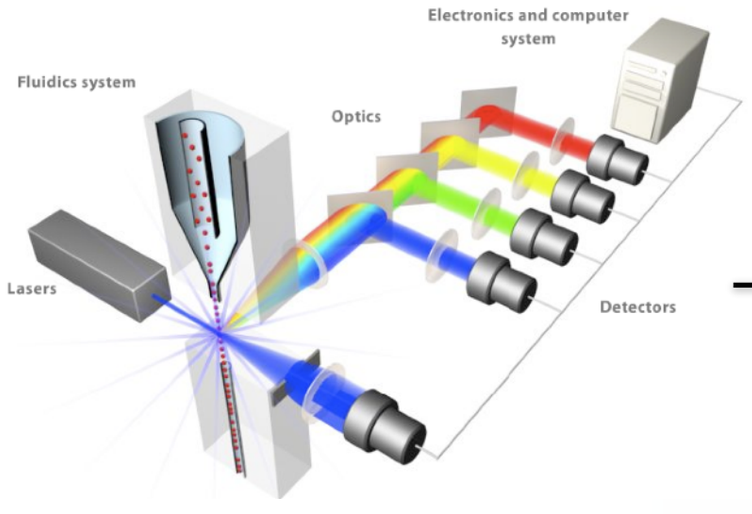
Experimental Assay



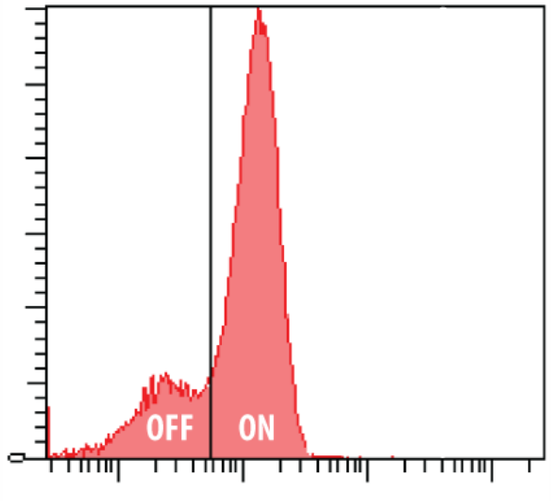




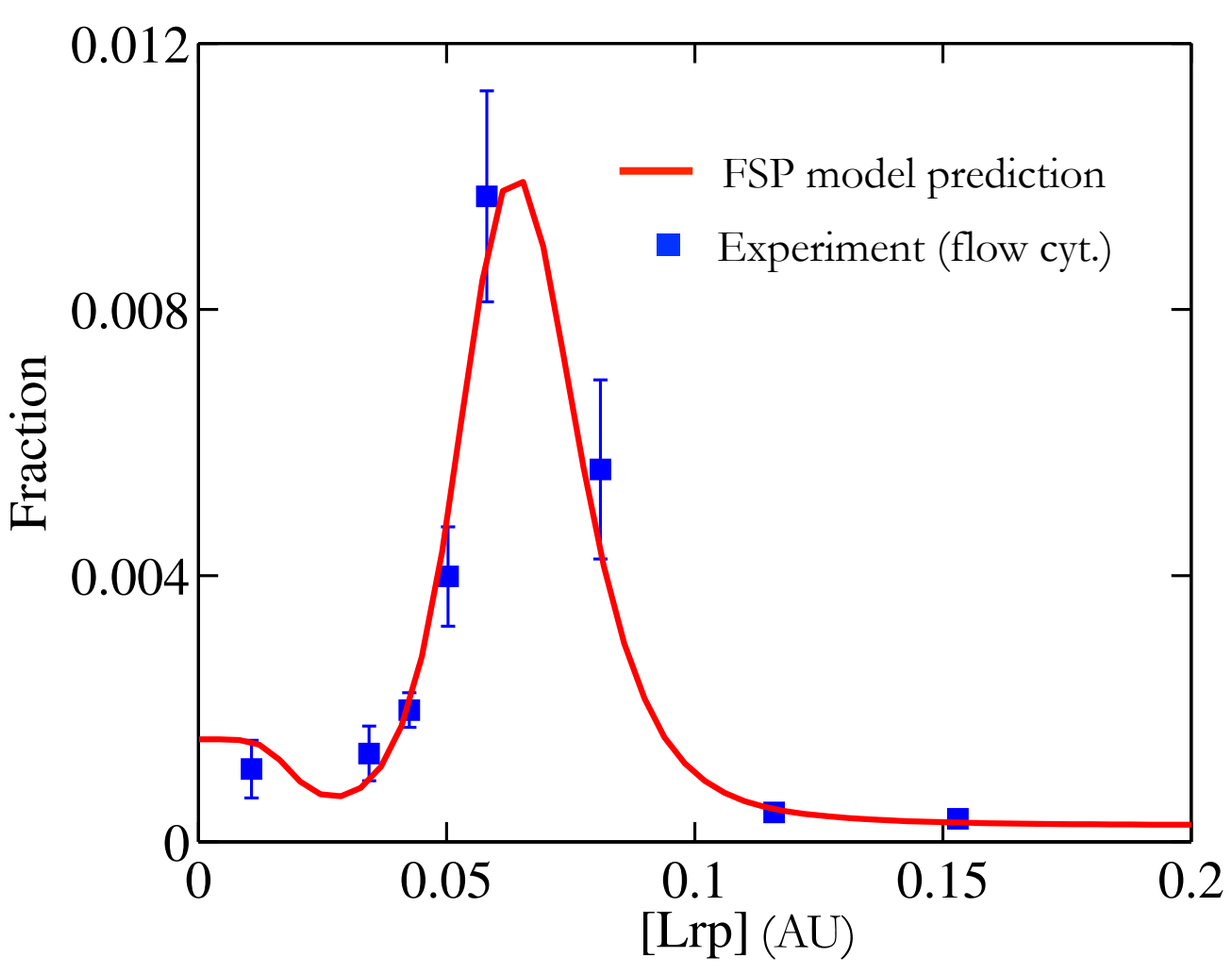
Flow Cytometry

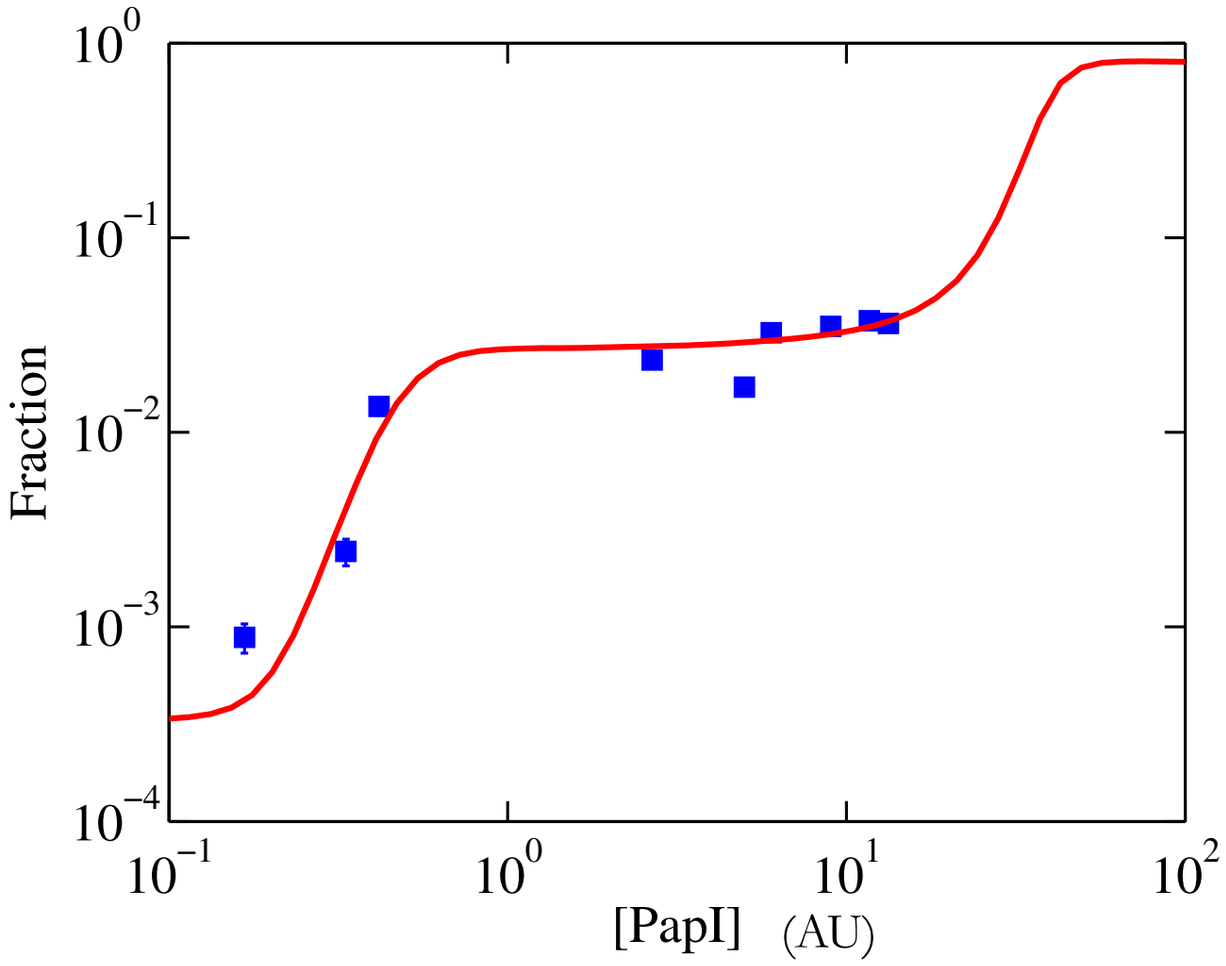


1 million cells per sample



Model Prediction





unpublished

Computing the Statistical Moments

Moment Closure Methods

Moment Computations

Moment Dynamics

$$\begin{aligned} \frac{dE[X]}{dt} &= S\ E[w(X)] \\ \frac{dE[XX^T]}{dt} &= SE[w(X)X^T] + E[Xw^T(X)]S^T + S\ diag(E[w(X)])\ S^T \end{aligned}$$

- Affine propensity. Closed moment equations.
- Quadratic propensity. Not generally closed.
  - *Mass Fluctuation Kinetics* (Gomez-Uribe, Verghese)
  - *Derivative Matching* (Singh, Hespanha)

Mass Fluctuation Kinetics (MFK) Moment Closure

Define $X_\Delta := X - \mathbb{E}(X)$, and $\Sigma := \mathbb{E}(X_\Delta X_\Delta^T)$.

We can write the propensity vector as follows

$$w(x) = \begin{bmatrix} w_1(x) \\ \vdots \\ w_M(x) \end{bmatrix} = \begin{bmatrix} a_1 + b_1^T x + x^T Q_1 x \\ \vdots \\ a_M + b_M^T x + x^T Q_M x \end{bmatrix}$$

where Q_k is symmetric.

Note that this form is general enough to fit propensities arising from all the elementary reactions.

Dynamics of the Mean

$$\frac{d\mathbb{E}[X]}{dt} = S\ \mathbb{E}[w(X)]$$

$$\begin{aligned}
\mathbb{E}(w_k(X)) &= \mathbb{E}(a_k + b_k^T X + X^T Q_k X) \\
&= a_k + b_k^T \mathbb{E}(X) + \mathbb{E}((X_\Delta + \mathbb{E}(X))^T Q_k (X_\Delta + \mathbb{E}(X))) \\
&= w_k(\mathbb{E}(X)) + \mathbb{E}(X_\Delta^T Q_k X_\Delta) \\
&= w_k(\mathbb{E}(X)) + tr(Q_k \Sigma)
\end{aligned}$$

Defining $z(\Sigma) = [tr(Q_1\Sigma) \ldots tr(Q_M\Sigma)]^T$, we have

$$\frac{d}{dt}\mathbb{E}(X) = Sw(\mathbb{E}(X)) + Sz(\Sigma).$$

Dynamics of the Covariance and MFK Closure

$$\frac{d\mathbb{E}[XX^T]}{dt} = S\mathbb{E}[w(X)X^T] + \mathbb{E}[Xw^T(X)]S^T + S\ diag(\mathbb{E}[w(X)])\ S^T$$

Moment closure assumption

The MFK approximation posits that $\forall i, j, k$:

$$\mathbb{E}\left[(X_i - \mathbb{E}(X_i))(X_j - \mathbb{E}(X_j))(X_k - \mathbb{E}(X_k))\right] \approx 0.$$

This enables the replacement of 3rd order moments with lower moments

The system is said to be closed and can be solved

$$\dot{\Sigma} = SJ_w(\mathbb{E}(X)) \cdot \Sigma + \Sigma \cdot J_w^T(\mathbb{E}(X))S^T + S\ diag(w(\mathbb{E}(X)))S^T + S\ diag(z(\Sigma))S^T$$

where J_w is the Jacobian matrix of $w(\cdot)$.

Other Moment Closure Methods

Another moment closure method is Derivative-Matching.

Moment closure assumption

The Derivative-Matching approximation posits that:

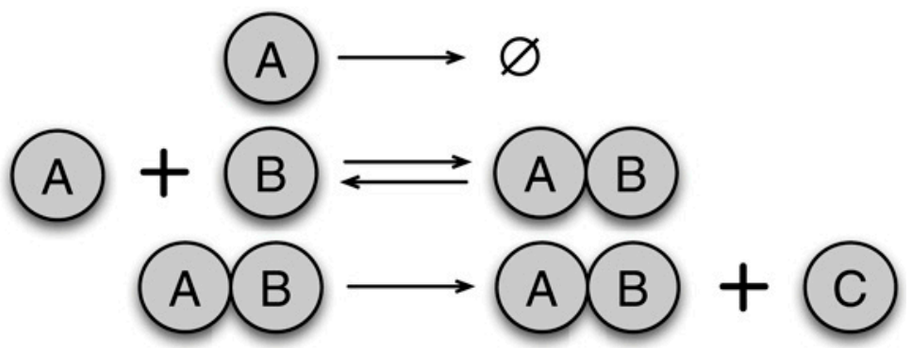
$$\mathbb{E}(X_i X_j X_k) = \frac{\mathbb{E}(X_i X_j)\mathbb{E}(X_j X_k)\mathbb{E}(X_i X_k)}{\mathbb{E}(X_i)\mathbb{E}(X_j)\mathbb{E}(X_k)}, \qquad \forall i, j, k.$$

This is consistent with a *Lognormal* moment closure.

Moment-Based Inference

Moment-Based Inference

A simplified model of gene expression



$$A \xrightarrow{c_1} \emptyset$$

$$A + B \underset{c_3}{\overset{c_2}{\rightleftharpoons}} AB$$

$$AB \xrightarrow{c_4} AB + C,$$

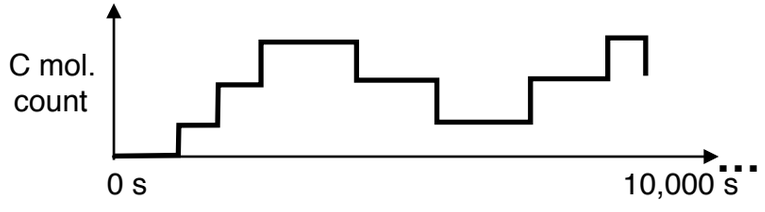
A is transcription factor
B is a gene
C is protein product

Parameter	c_1	c_2	c_3	c_4
γ_j	$1.500 \cdot 10^{-2}$ s^{-1}	$8.000 \cdot 10^{-4}$ s^{-1}	$1.000 \cdot 10^{-3}$ s^{-1}	$4.000 \cdot 10^{-1}$ s^{-1}

Initial conditions

A=50 molecules
B=1 (gene is inactive)
AB=0 molecules
C=0 molecules

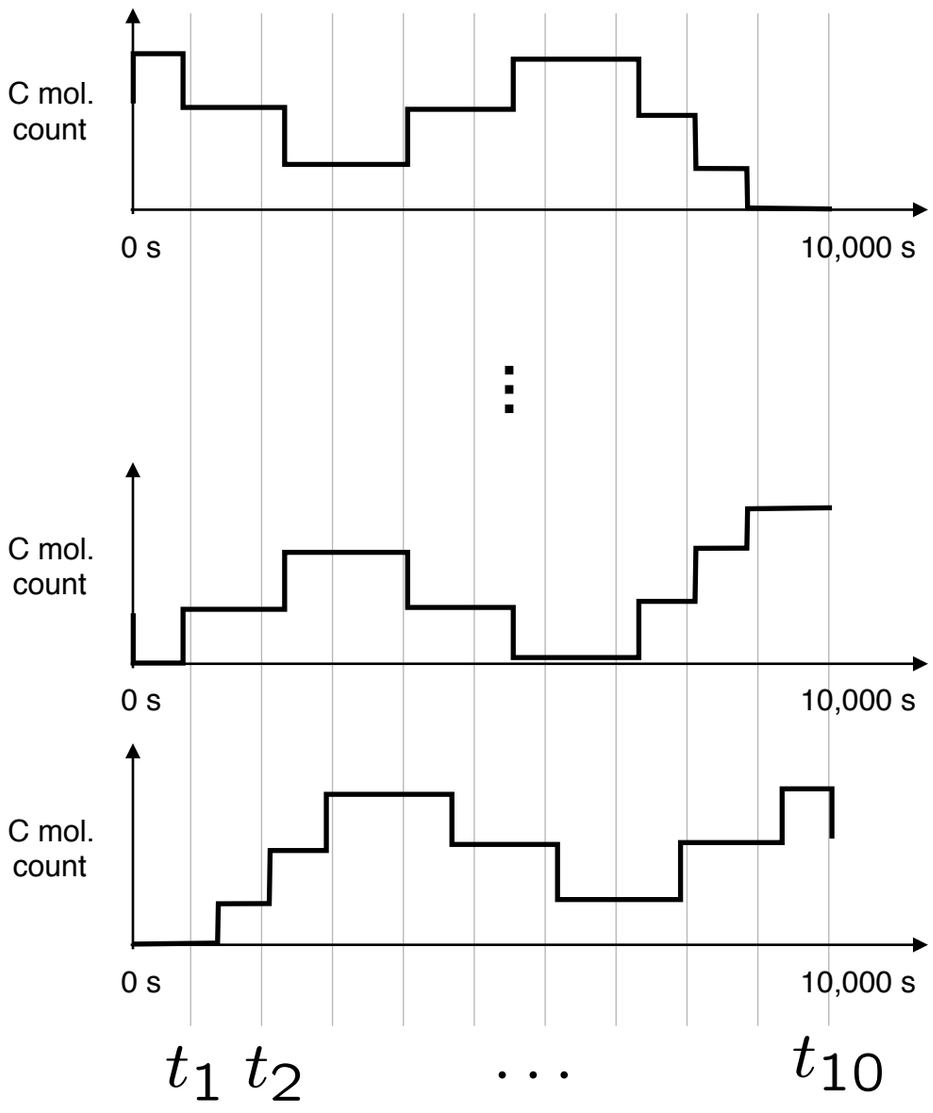
20,000 sample paths of length T=10,000s (generated using SSA algorithm)



Example adapted from *Zechner et al., PNAS, 2012*

Moment Based Inference: Data Model

Data: Trajectories samples over time



$\{x_i(t_l)\}_{i=1}^{M} \quad l = 1, \ldots, 10$

$M = 20,000$ samples

Data modeling and uncertainty

kth-order moment estimates:

$$\hat{\mu}_k(t_l) = \begin{cases} \frac{1}{M}\sum_{i=1}^{M} x_i(t_l) & k = 1 \\ \frac{1}{M}\sum_{i=1}^{M} (x_i(t_l) - \hat{\mu}_1(t_l))^k & k > 1. \end{cases}$$

CLT: For large M, moment estimates are normally distributed: $\hat{\mu}_k(t_l) \sim \mathcal{N}(\mu_k(t_l), \sigma_k^2(t_l))$

Estimator variance $\sigma_k^2(t_l)$ can be estimated:

Variance of Mean estimator:

$$\sigma_1^2(t_l) = \frac{1}{M}\hat{\mu}_2^2(t_l)$$

Variance of variance estimator:

$$\sigma_2^2(t_l) = \frac{1}{M}\left(\hat{\mu}_4(t_l) - \frac{M-3}{M-1}\hat{\mu}_2^2(t_l)\right)$$

Moments Model

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^1_A = c_3 \cdot \mu^1_D - c_2 \cdot \mu^2_{AB} - c_1 \cdot \mu^1_A$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^1_B = c_3 \cdot \mu^1_D - c_2 \cdot \mu^2_{AB}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^1_D = c_2 \cdot \mu^2_{AB} - c_3 \cdot \mu^1_D$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^1_C = c_4 \cdot \mu^1_D$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_A = c_1 \cdot \mu^1_A - 2 \cdot c_1 \cdot \mu^2_A + c_2 \cdot \mu^2_{AB} - 2 \cdot c_2 \cdot \boxed{\mu^3_{A2B}} + c_3 \cdot \mu^1_D + 2 \cdot c_3 \cdot \mu^2_{AD}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{AB} = c_3 \cdot \mu^1_D - c_2 \cdot \boxed{\mu^3_{AB2}} - c_2 \cdot \boxed{\mu^3_{A2B}} + c_3 \cdot \mu^2_{AD} + c_3 \cdot \mu^2_{BD} - (c_1 - c_2) \cdot \mu^2_{AB}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{AD} = c_2 \cdot \boxed{\mu^3_{A2B}} - c_2 \cdot \mu^2_{AB} - c_2 \cdot \boxed{\mu^3_{ABD}} - c_3 \cdot \mu^1_D + c_3 \cdot \mu^2_D - (c_1 + c_3) \cdot \mu^2_{AD}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{AC} = c_3 \cdot \mu^2_{DC} - c_2 \cdot \boxed{\mu^3_{ABC}} - c_1 \cdot \mu^2_{AC} + c_4 \cdot \mu^2_{AD}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_B = c_2 \cdot \mu^2_{AB} - 2 \cdot c_2 \cdot \boxed{\mu^3_{AB2}} + c_3 \cdot \mu^1_D + 2 \cdot c_3 \cdot \mu^2_{BD}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{BD} = c_2 \cdot \boxed{\mu^3_{AB2}} - c_2 \cdot \mu^2_{AB} - c_2 \cdot \boxed{\mu^3_{ABD}} - c_3 \cdot \mu^1_D + c_3 \cdot \mu^2_D - c_3 \cdot \mu^2_{BD}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{BC} = c_3 \cdot \mu^2_{DC} - c_2 \cdot \boxed{\mu^3_{ABC}} + c_4 \cdot \mu^2_{BD}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_D = c_2 \cdot \mu^2_{AB} + 2 \cdot c_2 \cdot \boxed{\mu^3_{ABD}} + c_3 \cdot \mu^1_D - 2 \cdot c_3 \cdot \mu^2_D$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_{DC} = c_2 \cdot \boxed{\mu^3_{ABC}} - c_3 \cdot \mu^2_{DC} + c_4 \cdot \mu^2_D$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\mu^2_C = c_4 \cdot \mu^1_D + 2 \cdot c_4 \cdot \mu^2_{DC}$$

Notation:

$$\mu_A := \mathbb{E}[X_A]$$
$$\mu^2_{AB} := \mathbb{E}[X_A X_B] \quad \mu^2_A := \mathbb{E}[X^2_A]$$
$$\mu^3_{ABC} := \mathbb{E}[X_A X_B X_C]$$
$$\vdots$$

Moments are not closed!

Closing the System of Moments

A closed system is obtained by replacing the 3rd-order cumulants by 0.

This is equivalent to replacing the 3rd-order by functions of the lower order moments:

$$\mu^3_{ABD} = \mu^1_A \cdot \mu^2_{BD} + \mu^1_B \cdot \mu^2_{AD} + \mu^1_D \cdot \mu^2_{AB} - 2 \cdot \mu^1_A \cdot \mu^1_B \cdot \mu^1_D,$$
$$\mu^3_{A2B} = -2 \cdot \mu^1_B \cdot {\mu^1_A}^2 + 2 \cdot \mu^2_{AB} \cdot \mu^1_A + \mu^1_B \cdot \mu^2_A,$$
$$\mu^3_{ABC} = \mu^1_A \cdot \mu^2_{BC} + \mu^1_B \cdot \mu^2_{AC} + \mu^1_C \cdot \mu^2_{AB} - 2 \cdot \mu^1_A \cdot \mu^1_B \cdot \mu^1_C,$$
$$\mu^3_{AB2} = -2 \cdot \mu^1_A \cdot {\mu^1_B}^2 + 2 \cdot \mu^2_{AB} \cdot \mu^1_B + \mu^1_A \cdot \mu^2_B$$

Moment Model:

$$\frac{d}{dt}\tilde{\mu} = A(\theta)\tilde{\mu} + B(\theta) f(\tilde{\mu})$$

$\tilde{\mu}$ is the vector of 1st- and 2nd-order moments

Moment-Based Inference

$$p(\theta|\hat{\mu}_1, \hat{\mu}_2) = \frac{1}{K} \prod_{k=1}^{2} \prod_{l=1}^{10} p(\hat{\mu}_k(t_l)|\theta) \cdot p(\theta)$$

Used **Metropolis-Hastings** MCMC:

For each parameter in θ, a log-normal proposal distribution was used

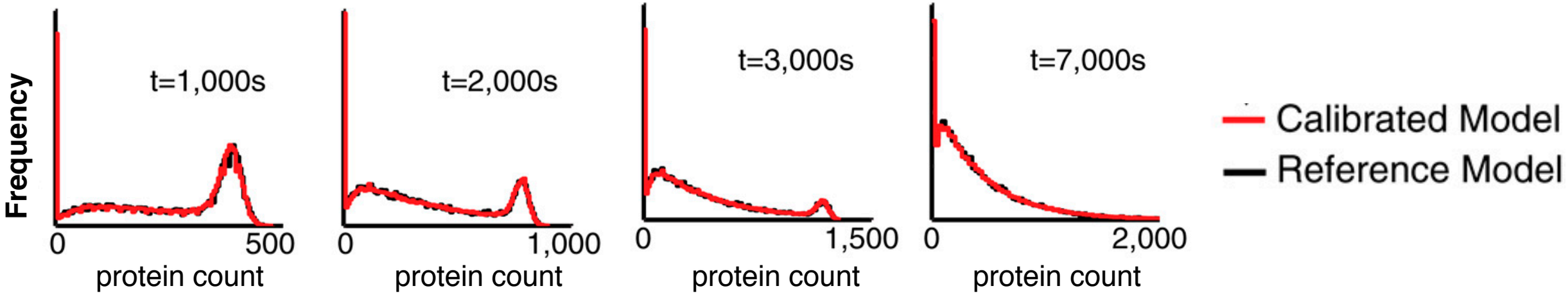
$$q(\theta^{\text{new}}|\theta^{\text{old}}) = \prod_{j=1}^{4} q(\theta_j^{\text{new}}|\theta_j^{\text{old}}) \qquad q(\theta_j^{\text{new}}|\theta_j^{\text{old}}) = \mathcal{LN}(\ln \theta_j^{\text{old}}, v_j^2)$$

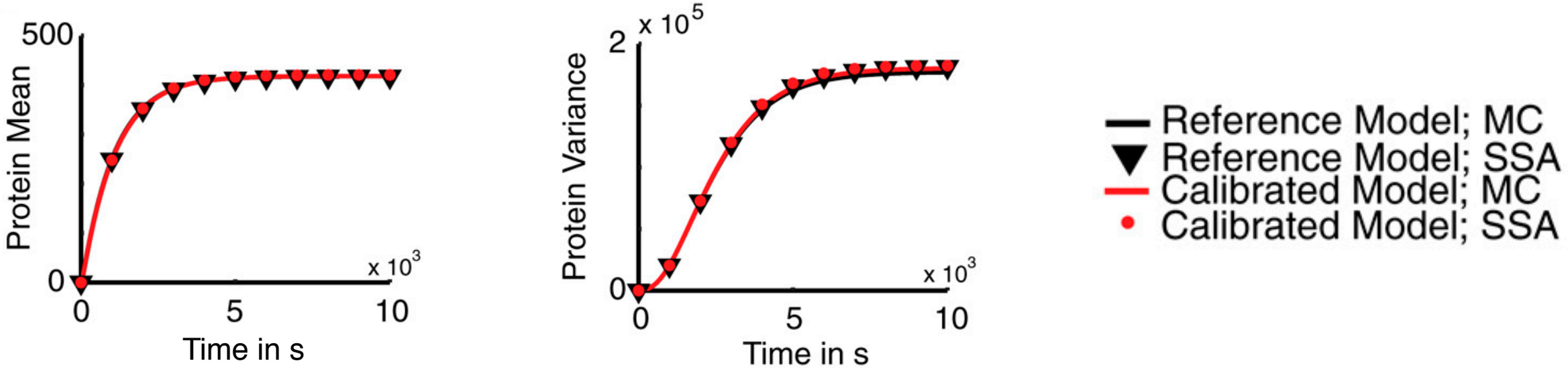
Inferred Parameter Vector

Parameter	c_1	c_2	c_3	c_4
θ_j	$1.500 \cdot 10^{-2}$	$8.000 \cdot 10^{-4}$	$1.000 \cdot 10^{-3}$	$4.000 \cdot 10^{-1}$
$\theta_{j,MAP}$	$1.380 \cdot 10^{-2}$	$7.050 \cdot 10^{-4}$	$9.865 \cdot 10^{-4}$	$3.988 \cdot 10^{-1}$
	s^{-1}	s^{-1}	s^{-1}	s^{-1}

$$\theta_{MAP} = \arg\max_{\theta} \; p(\theta|\hat{\mu}_1, \hat{\mu}_2)$$

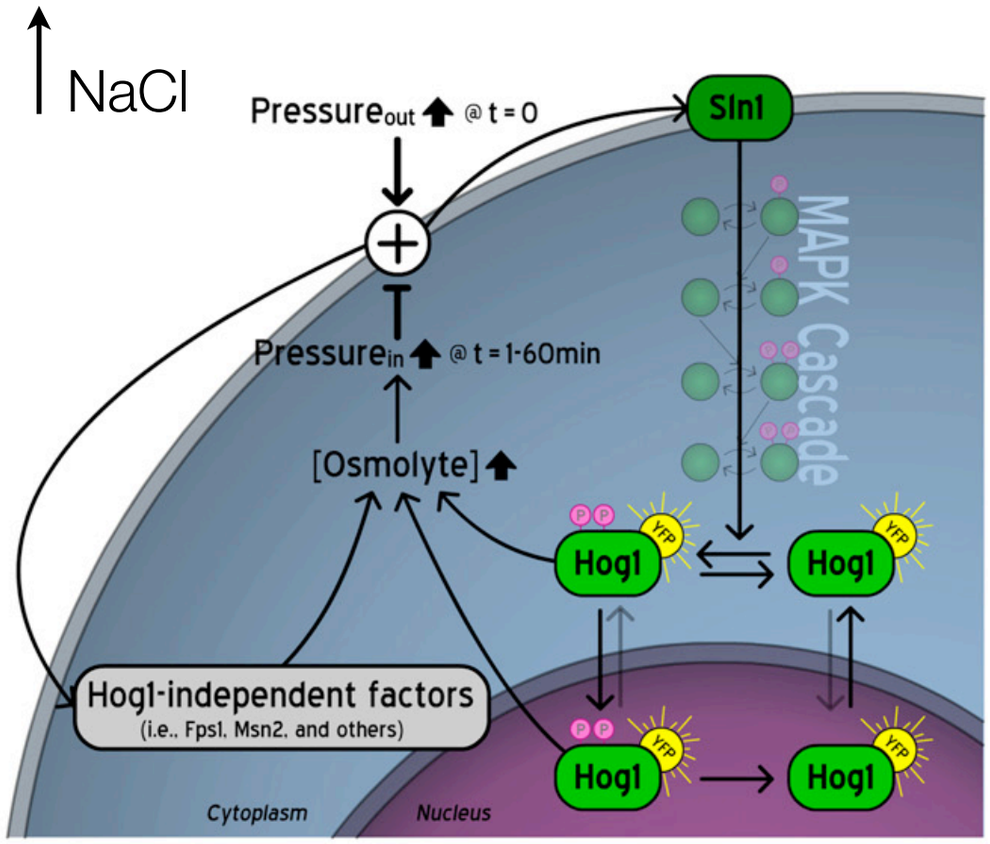
Calibrated Model vs. Reference Model





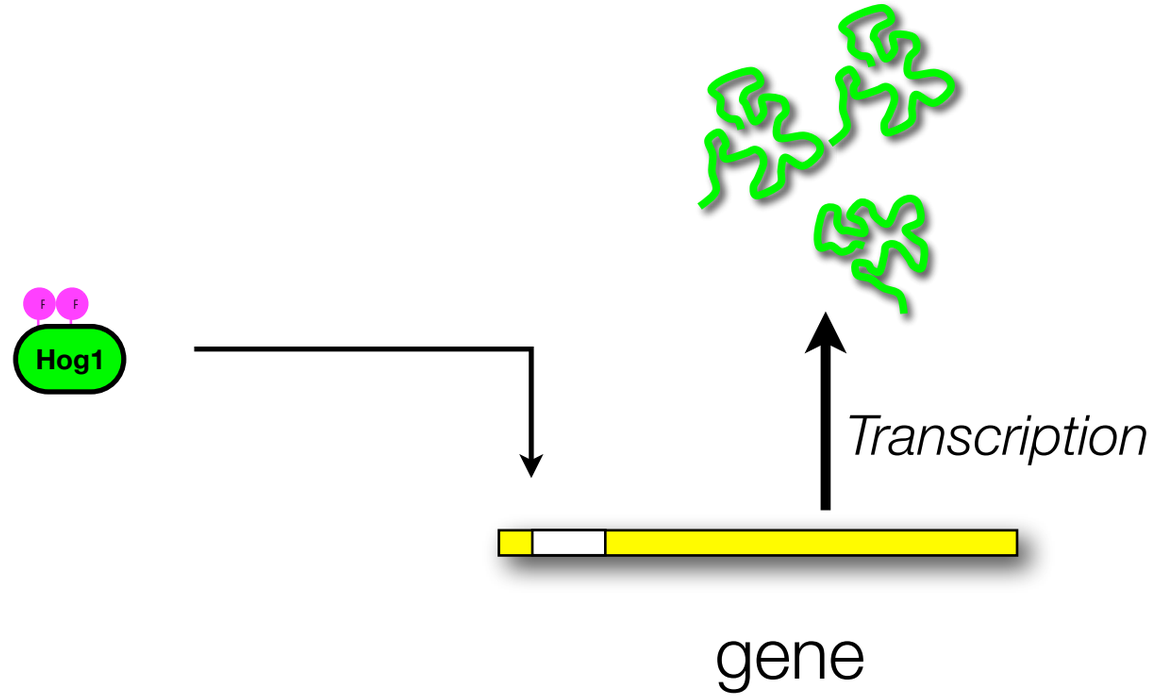
Density-Based Inference

Density-Based Estimation



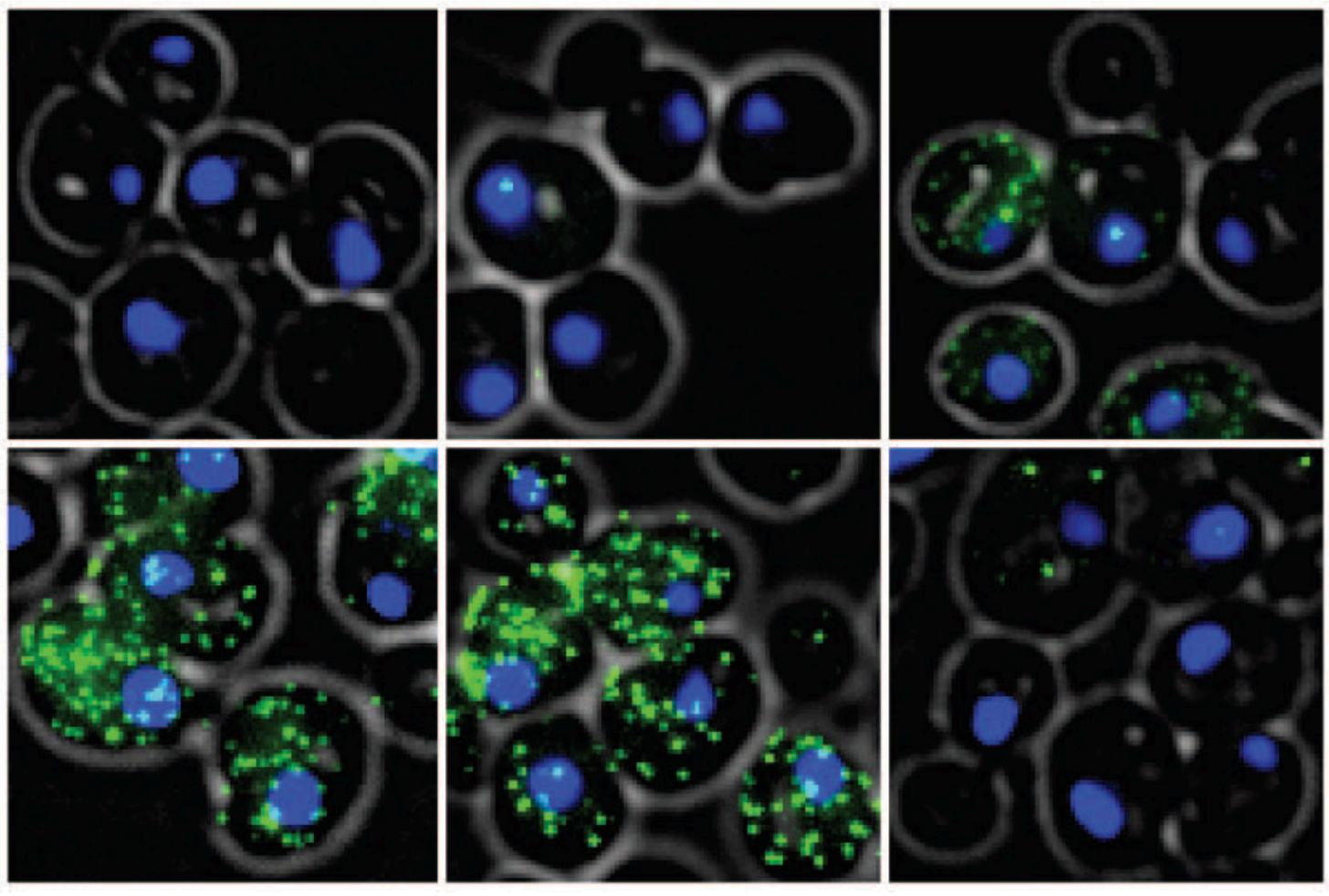
Muzzey et. al, Cell (2009)

Yeast Osmoregulation



Goal: Identify model of mRNA transcription

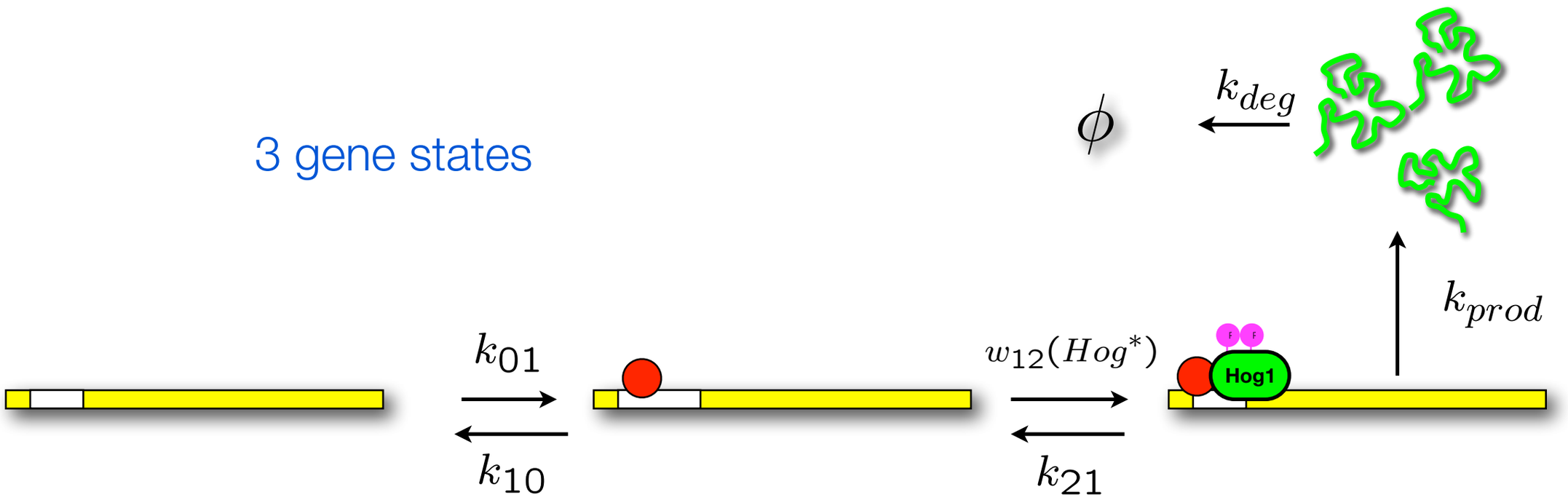
mRNA Measurements



- *mRNA* copy numbers are measured using mRNA FISH method
- histograms give snapshots of pdf.

Identifying Gene Parameters from mRNA Data

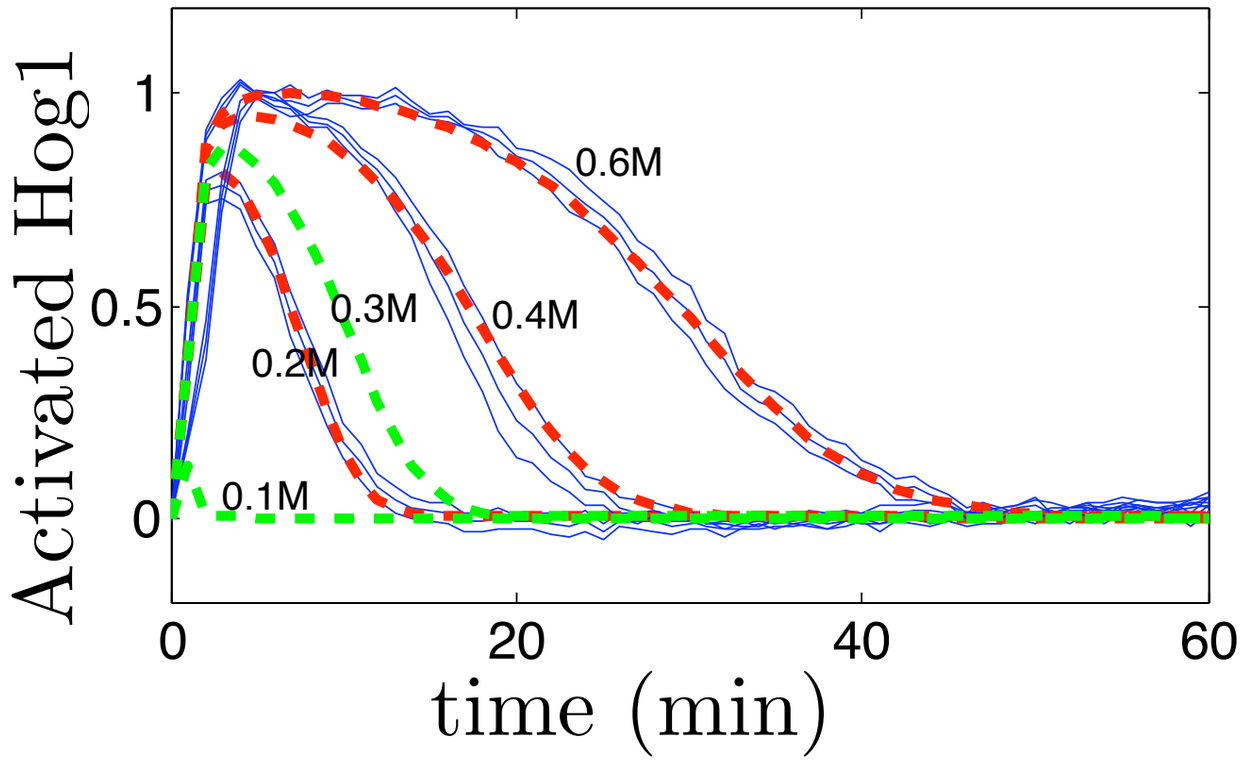
3 gene states



$$w_{12}(Hog1^\star) = k_{12}[Hog1^\star]^\eta$$

7 parameters are unknown

Modeling Hog1* Nuclear Concentration



$$Hog1(t) = \left(1 - e^{-r_1 t}\right) e^{-\frac{\alpha t}{[NaCl]}}$$

$$Hog1^*(t) = \frac{Hog1(t)}{1 + Hog1(t)/M}$$

Density-Based Estimation

Given N numbered cells.

Suppose the nth cell was measured at time t_n to have exactly m_n copies of mRNA.

The likelihood that the data from the N cells came from the model with parameter θ is given by:

$$L(D|\theta) = \prod_{n=1}^{N} p(m_n|\theta, t_n)$$

$p(m_n|\theta, t_n)$ is given by the chemical master equation

$$\theta_{ML} = \arg\max_{\theta} \log L(D|\theta)$$

Finite State Projection

$P_{i,m} := P(\text{state} = S_i, \text{mRNA} = m).$

Enumerate all the possible states into the vector:

$$\begin{bmatrix} \mathbf{P}_0 \\ \\ \mathbf{P}_1 \\ \\ \vdots \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathbf{P}_{1,0} \\ \mathbf{P}_{2,0} \\ \vdots \\ \mathbf{P}_{N,0} \end{bmatrix} \\ \begin{bmatrix} \mathbf{P}_{1,1} \\ \mathbf{P}_{2,1} \\ \vdots \\ \mathbf{P}_{N,1} \end{bmatrix} \\ \vdots \end{bmatrix},$$

Since the gene can take one of three possible states, $N = 3$

The maximum number of mRNA we expect to see is 150

Truncate $P \to P_{FSP}$

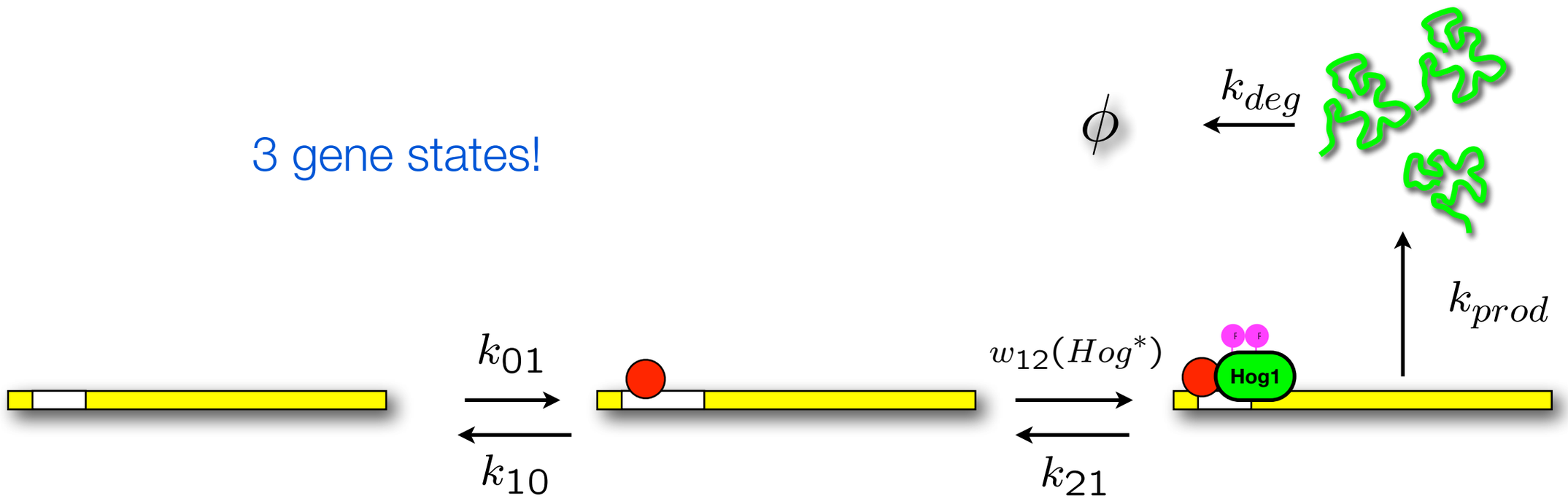
Solve $\dot{P}_{FSP} = A(\theta) P_{FSP}$

$\|P_{FSP}(t) - P(t)\|_1 < 10^{-6}$

$0 \leq t \leq 80\text{min}$

Inferring Gene Parameters from mRNA Data

3 gene states!



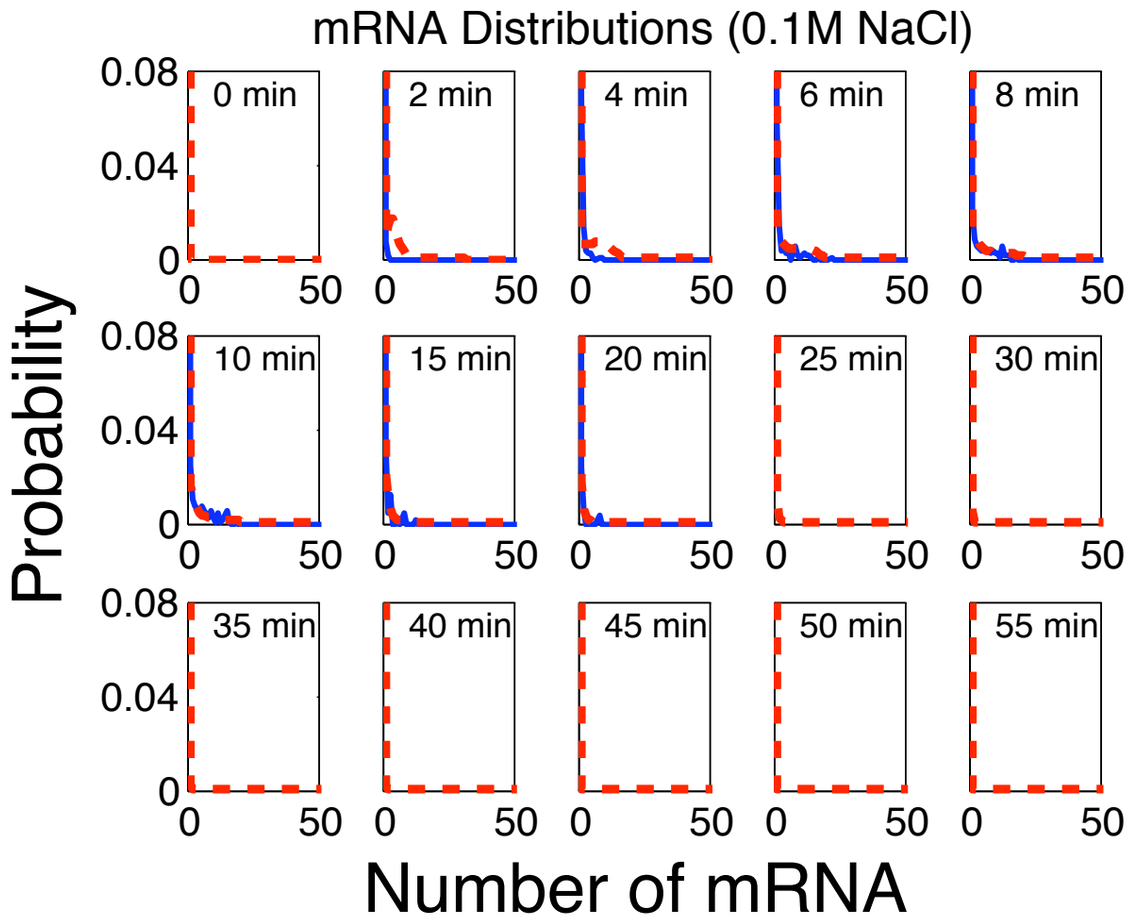
$$k_{01} = 1.3063 \times 10^{-3}\mathrm{s}^{-1} \qquad k_{21} = 3.7763 \times 10^{-3}\mathrm{s}^{-1}$$

$$k_{10} = 6.4891 \times 10^{-4}\mathrm{s}^{-1} \qquad k_{12} = 1.1691 \times 10^{9}\mathrm{Mol}^{-\eta}\mathrm{s}^{-1}$$

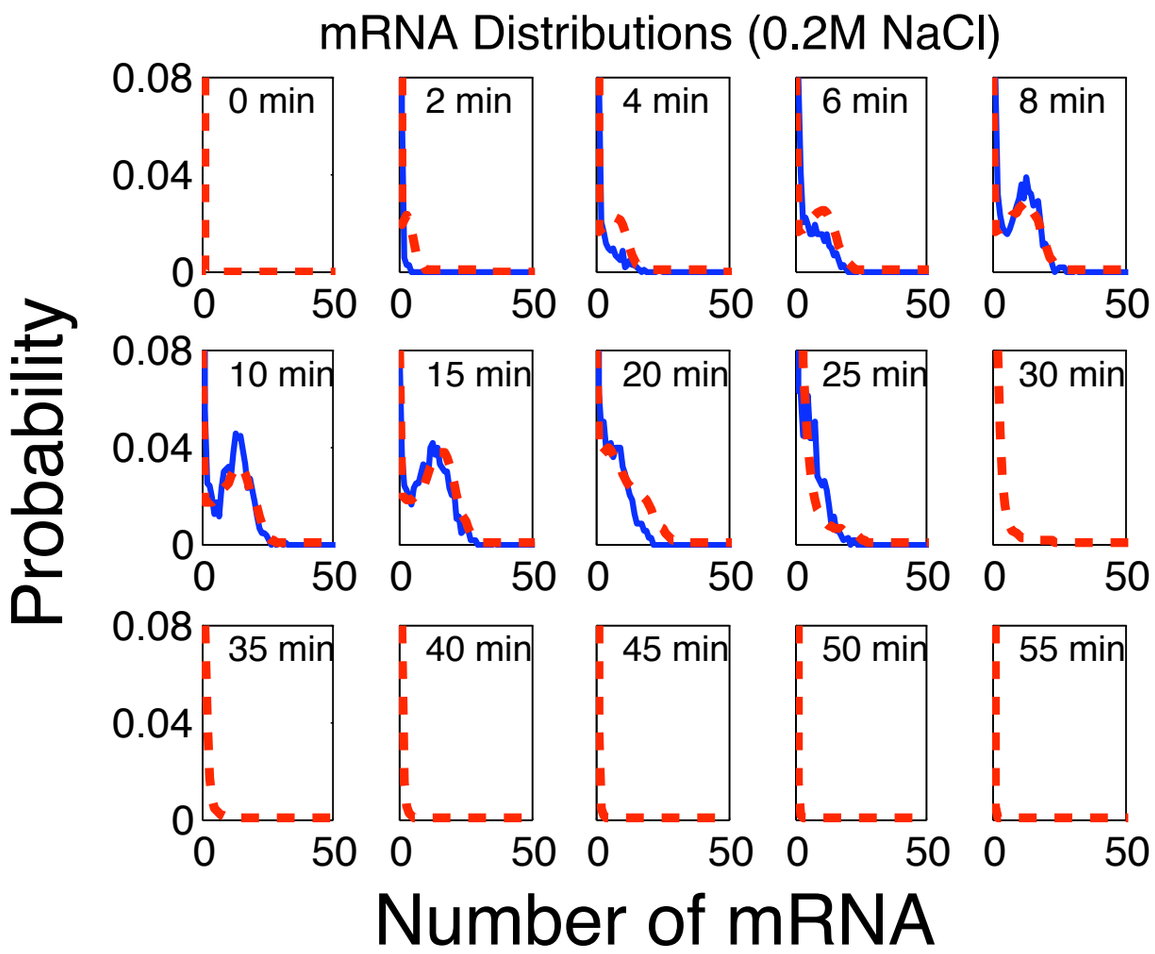
$$K_{deg} = 3.7282 \times 10^{-3}\mathrm{Mol}^{-1}\mathrm{s}^{-1} \qquad k_{prod} = 7.8818 \times 10^{-2}\mathrm{s}^{-1}$$

$$\eta = 2.2444$$

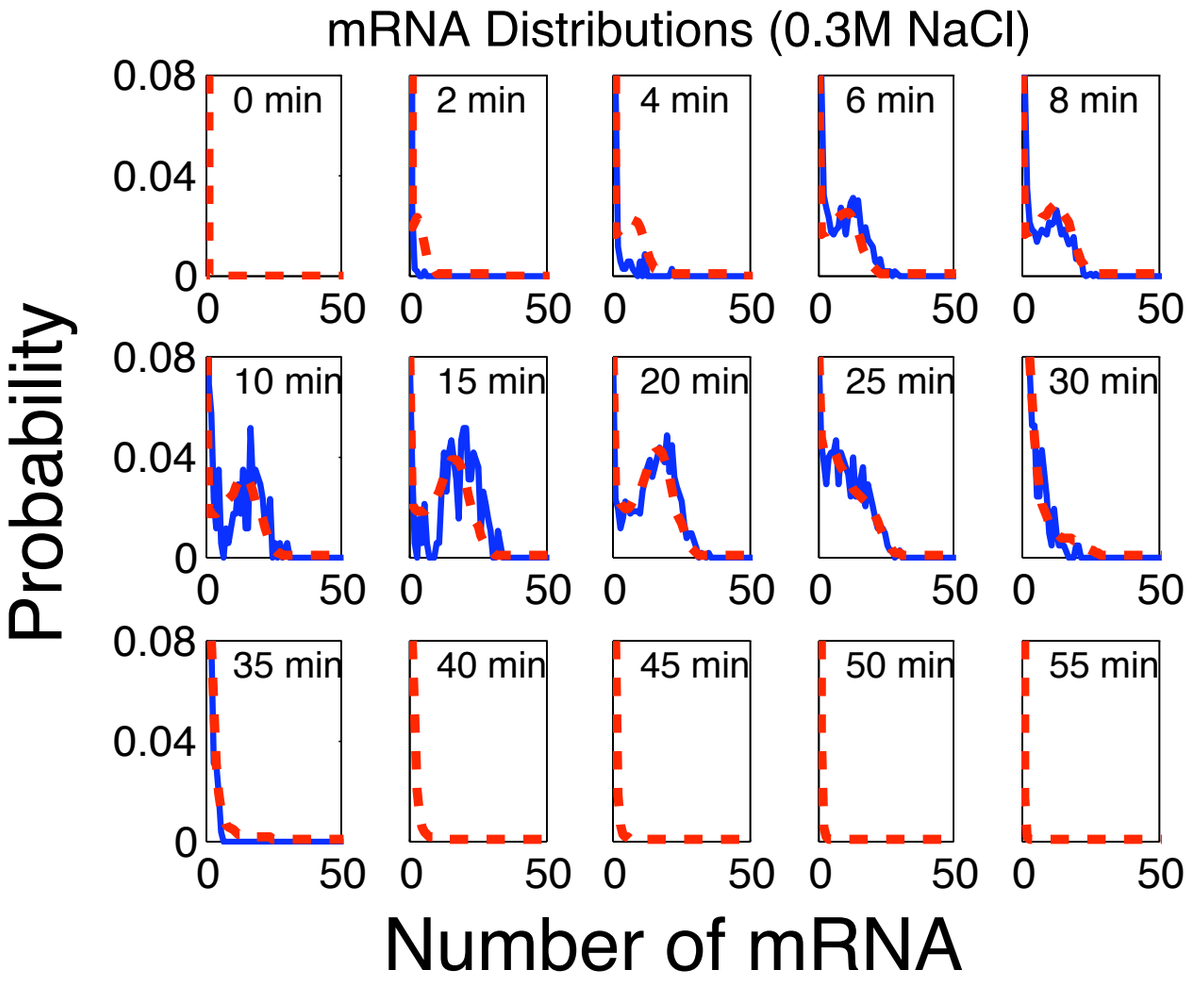
$$w_{12}(Hog1^\star) = k_{12}[Hog1^\star]^\eta$$



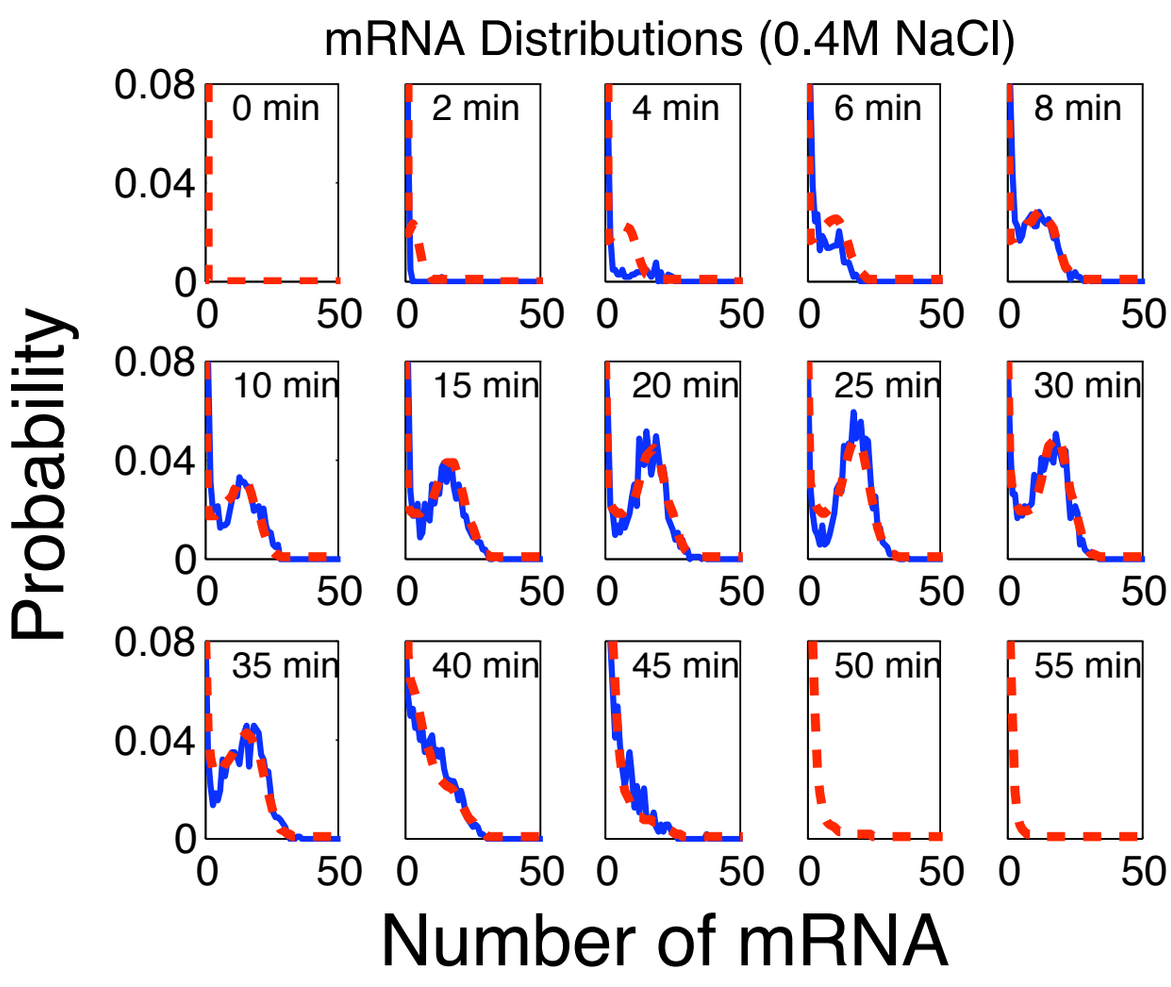
0.1M NaCl



0.2M NaCl



0.3M NaCl



0.4M NaCl

Directions

Computational Analysis of Stochastic Kinetics

Master Equation	Monte Carlo	SDE Approximation	Moment Dynamics	Hybrid
▸State Aggregation ▸Sparse computations ▸Basis expansion ▸Sensitivity Analysis ▸Inference	▸Speeding up SSA ▸Variance reduction ▸Sensitivity Analysis ▸Stochastic optimization ▸Inference	▸Linear Noise Approx. ▸Langevin equations ▸Time-scale separation ▸Spatial models ▸Inference	▸Better moment closure ▸Error bounds ▸Reduced state-space ▸Inference	▸Stoch/deterministic ▸SSA/FSP ▸Large count/low count partition ▸Inference

Application in Biology

Reverse Engineering

Noise & Function	Noise Exploitation	Noise Suppression
Role of Regulation	Effect on Fitness	Recurring Motifs

Synthetic Biology

Effect of Noise on Circuits	Design Architecture	Parameter Selection
Robustness	Noise transmission	Fundamental Limitations