The Chemical Master Equation in Gene Networks: Complexity and Approaches

Mustafa Khammash

Center for Control, Dynamical-Systems, and Computations
University of California at Santa Barbara

Outline

- Gene expression
  - An introduction
  - Stochasticity in gene expression
  - Randomness and variability
- Stochastic Chemical Kinetics
  - The Chemical Master Equation (CME)
  - Monte-Carlo simulations of the CME
- A new direct approach to solving the CME
  - The Finite-State Projection Method
  - Model reduction and aggregation
  - Examples

Gene Expression

- An Introduction
- Stochasticity in Gene Expression
- Randomness and variability

Why Are Stochastic Models Needed?

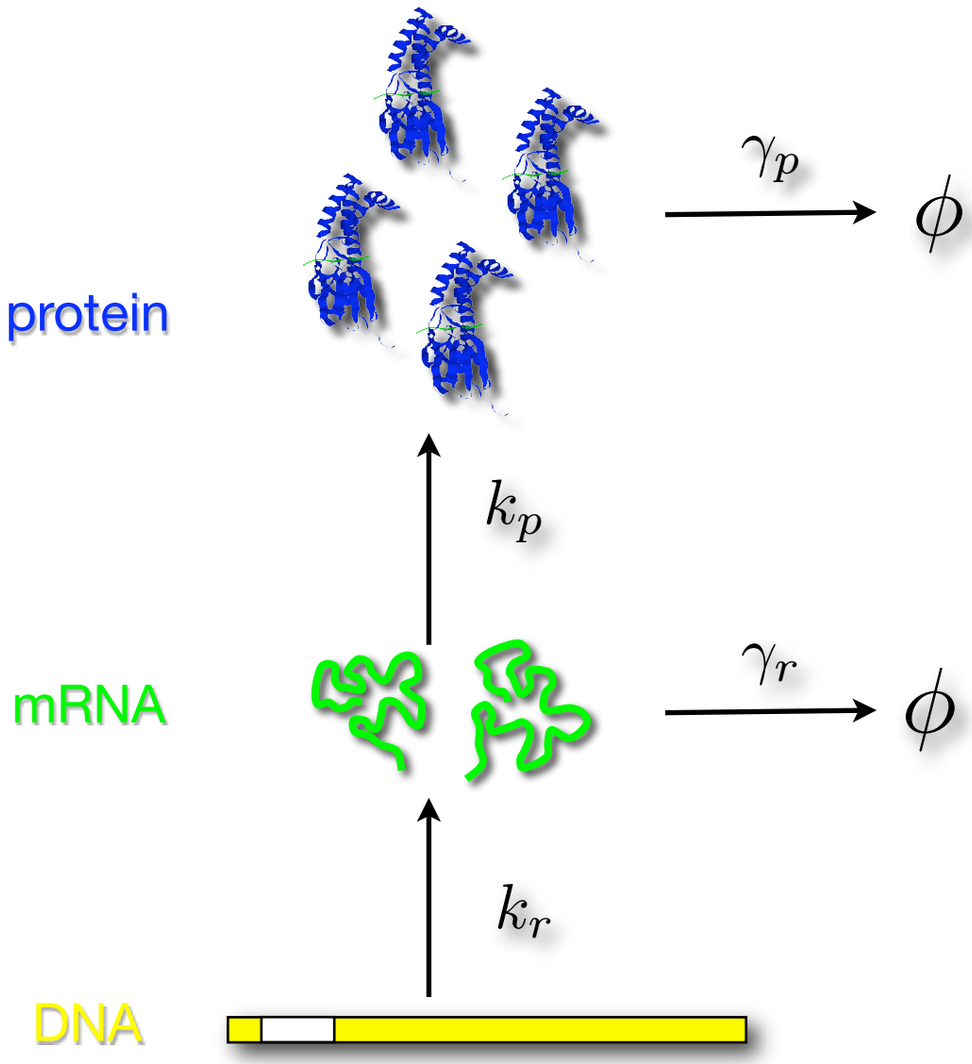
- Much of the mathematical modeling of gene networks represents gene expression deterministically

- Deterministic models describe macroscopic behavior; but many cellular constituents are present in small numbers

- Considerable experimental evidence indicates that significant stochastic fluctuations are present

- There are many examples when deterministic models are *not adequate*

Modeling Gene Expression



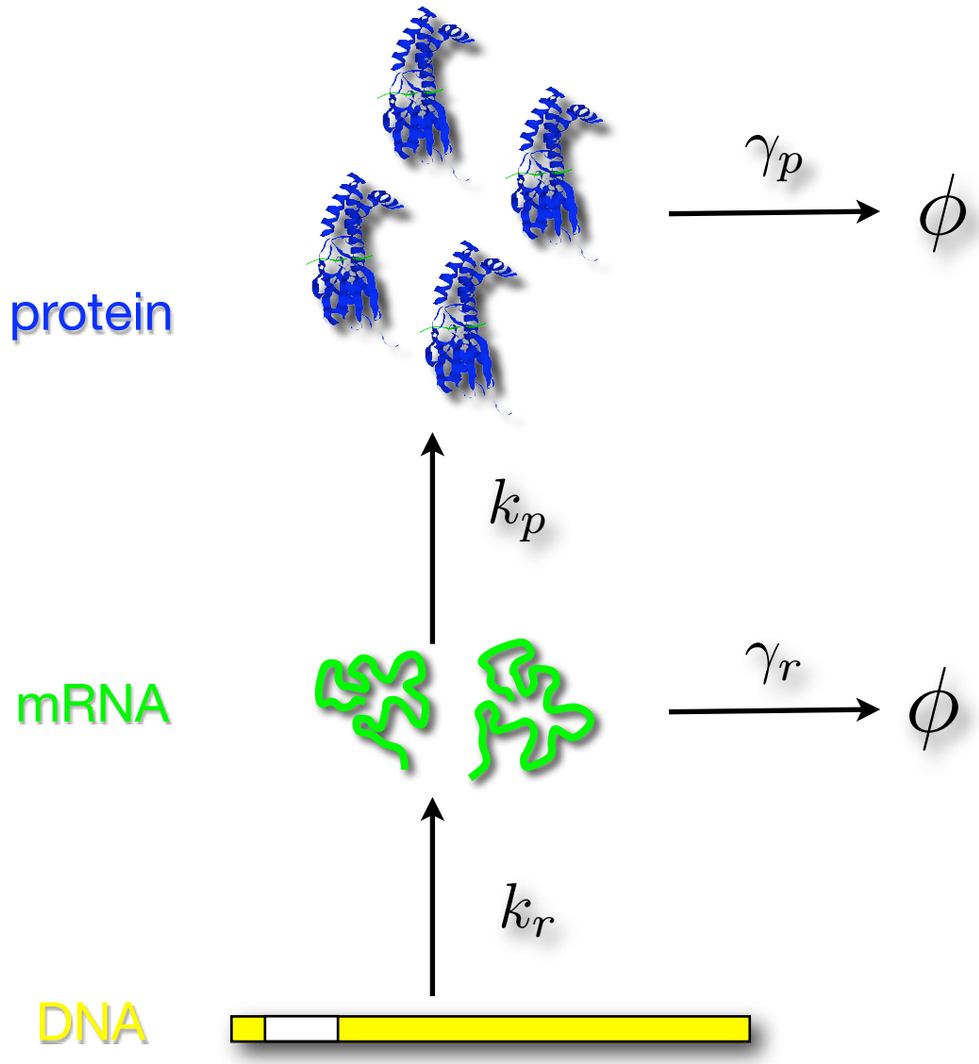
Deterministic model

$$\frac{d[mRNA]}{dt} = -\gamma_r[mRNA] + k_r$$
$$\frac{d[protein]}{dt} = -\gamma_p[protein] + k_p[mRNA]$$

Stochastic model

- Probability a single mRNA is transcribed in time dt is $k_r dt$.
- Probability a single mRNA is degraded in time dt is $(\#mRNA) \cdot \gamma_r dt$

Fluctuations at Small Copy Numbers



$$E(p) = \frac{k_r k_p}{\gamma_r \gamma_p}$$

$$\eta(p) = \frac{1}{\sqrt{E(p)}}\left(1 + \frac{kp}{\gamma_p + \gamma_r}\right)^{1/2} \quad \text{(protein)}$$

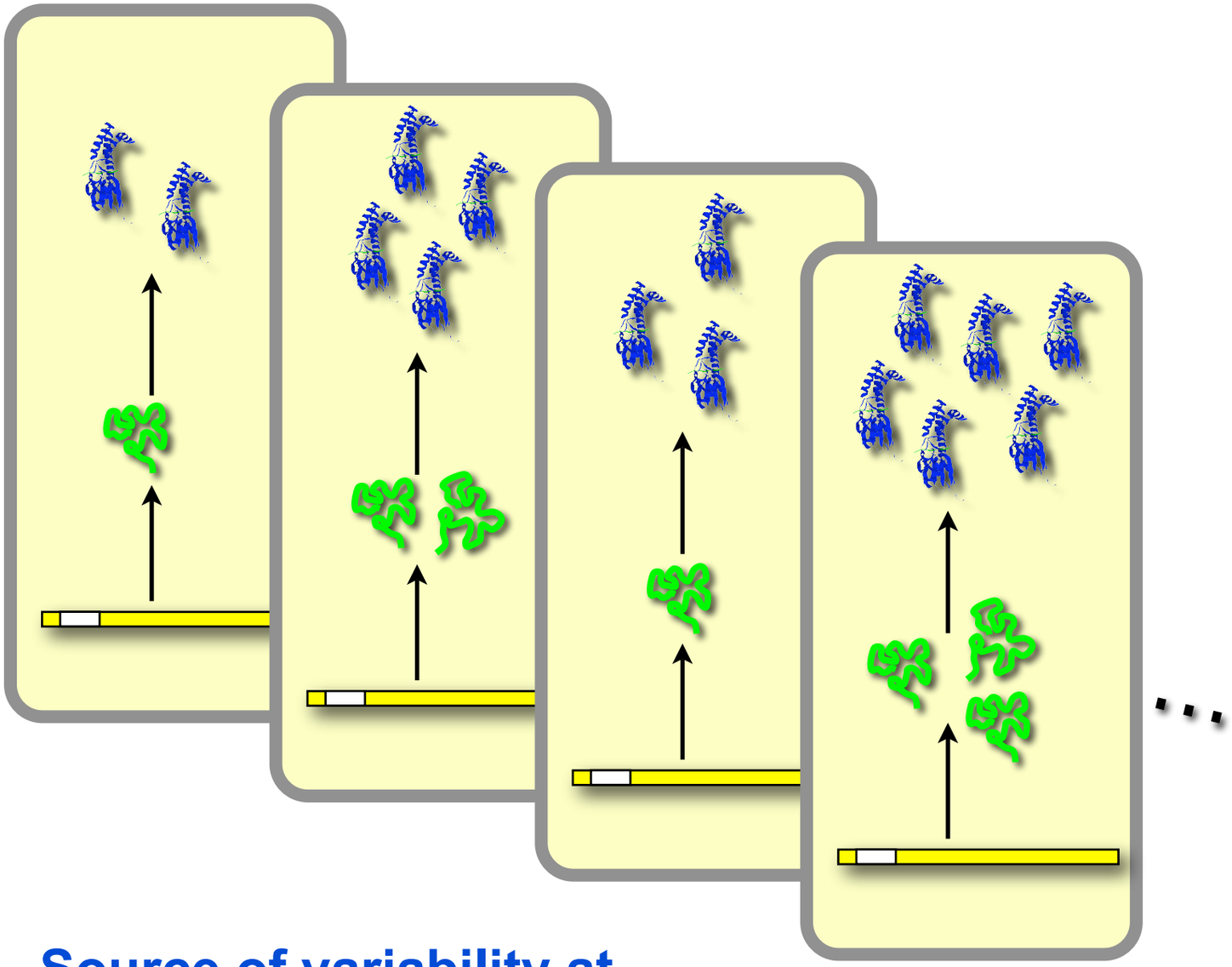
$$E(r) = \frac{k_r}{\gamma_r}$$

$$\eta(r) = \frac{1}{\sqrt{E(r)}} \quad \text{(mRNA)}$$

$$\text{Coefficient of variation}(\eta) = \frac{\text{standard deviation}}{\text{mean}}$$

- ★ Deterministic steady-state equals stochastic mean
- ★ Coefficient of variation goes as $1/\sqrt{\text{mean}}$
- ★ When mean is large, the coefficient of variation is (relatively) small

Intrinsic Variability in Gene Expression



Source of variability at cellular level….

- Small # of molecules
- Random events

“Intrinsic noise”

Impact of variability

- Noise propagates through the network
- Its amount depends on
  - # of molecules
  - stoichiometry
  - regulation
  - ...
- Sometimes it is suppressed; other times it is exploited
- Deterministic models are not adequate

Stochastic Influences on Phenotype

Fingerprints of identical twins

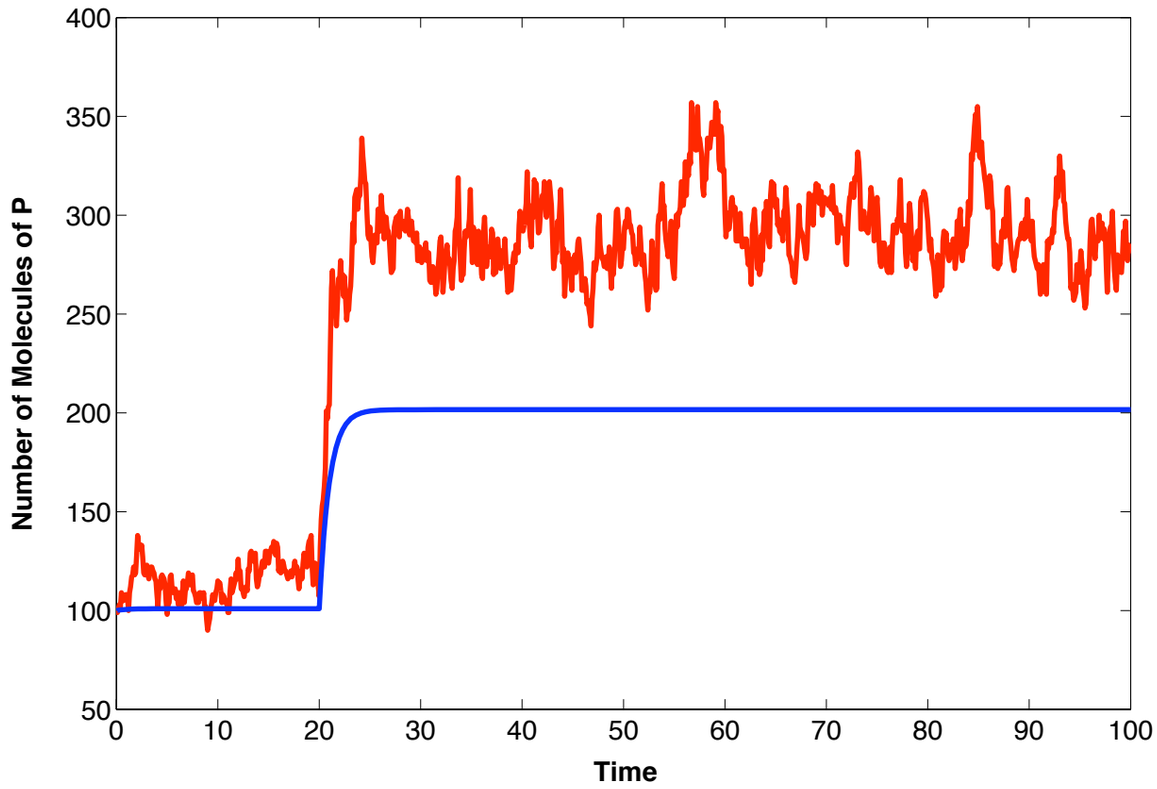
Cc, the first cloned cat and her genetic mother

J. Raser and E. O'Shea, Science, 1995.

Exploiting the Noise

$$\phi \underset{k_a S}{\overset{k}{\rightleftharpoons}} I \xrightarrow{k_p} P \xrightarrow{1} \phi$$

$$\phi \underset{k_d}{\overset{k_s}{\rightleftharpoons}} S$$

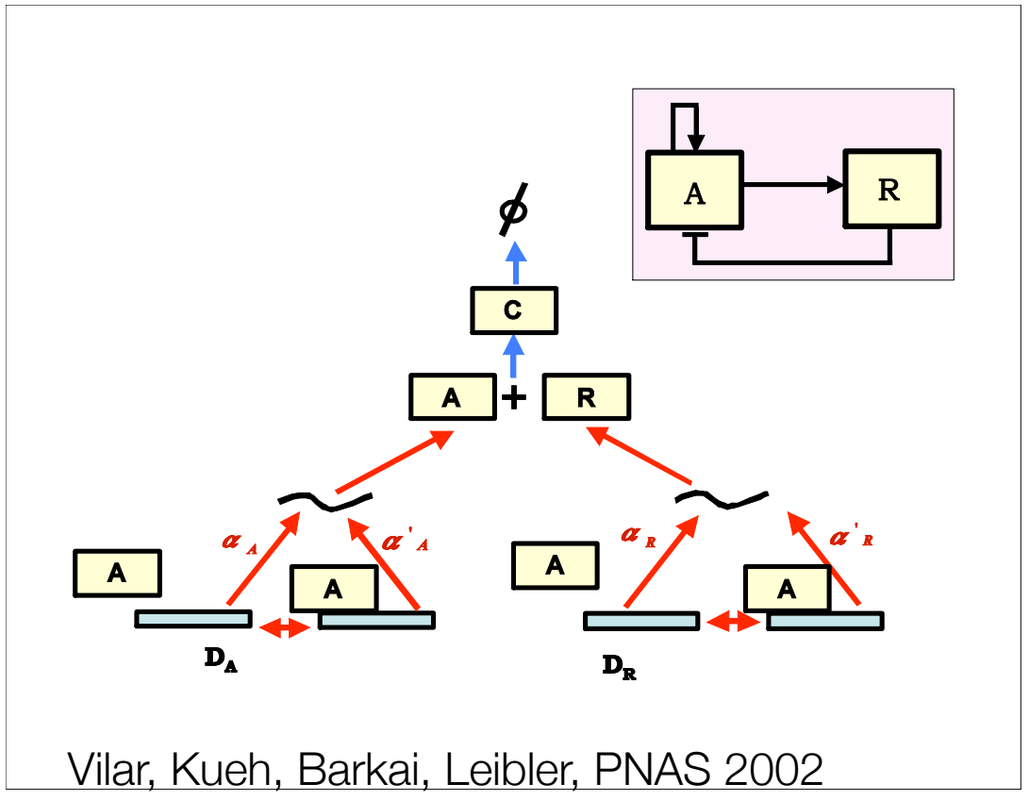


Johan Paulsson , Otto G. Berg , and Måns Ehrenberg, PNAS 2000

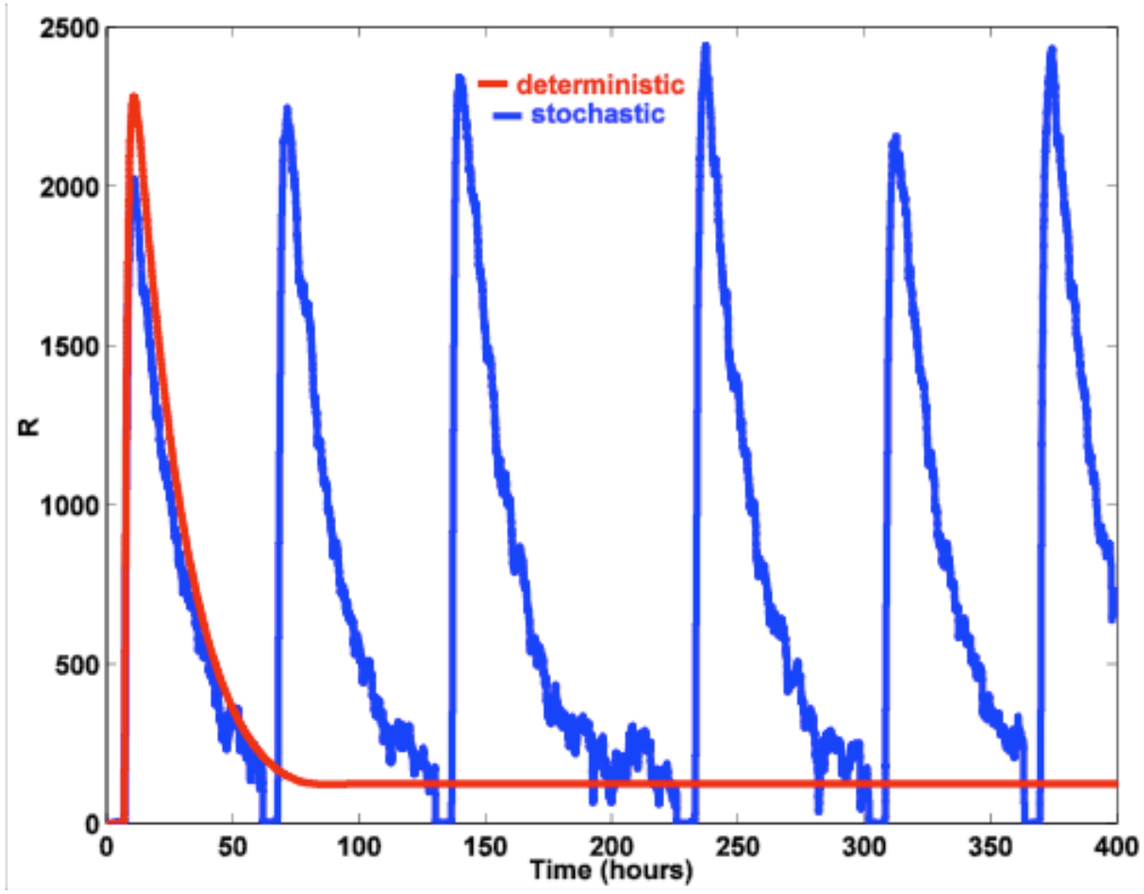
- Stochastic mean value different from deterministic steady state
- Noise *enhances* signal!

Noise Induced Oscillations

Circadian rhythm

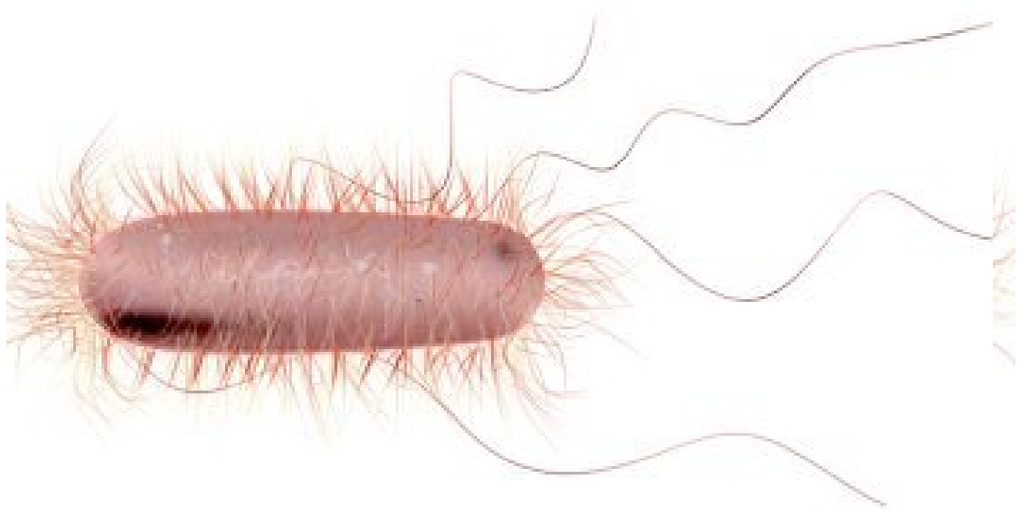


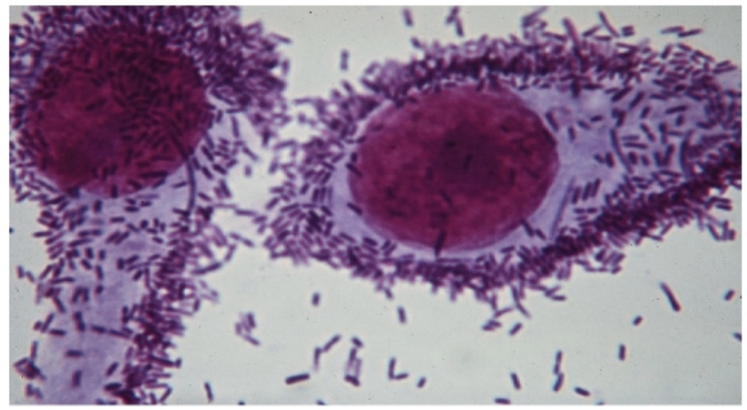
Vilar, Kueh, Barkai, Leibler, PNAS 2002



- Oscillations disappear from deterministic model after a small reduction in deg. of repressor
- (Coherence resonance) Regularity of noise induced oscillations can be manipulated by tuning the level of noise [El-Samad, Khammash]

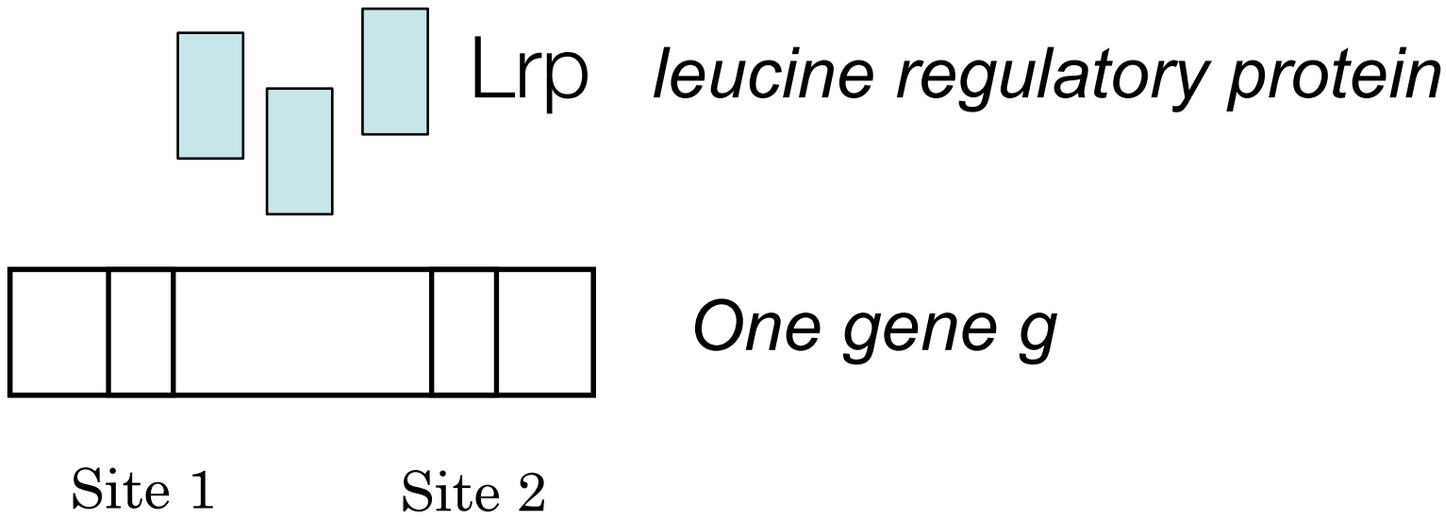
The Pap Pili Stochastic Switch



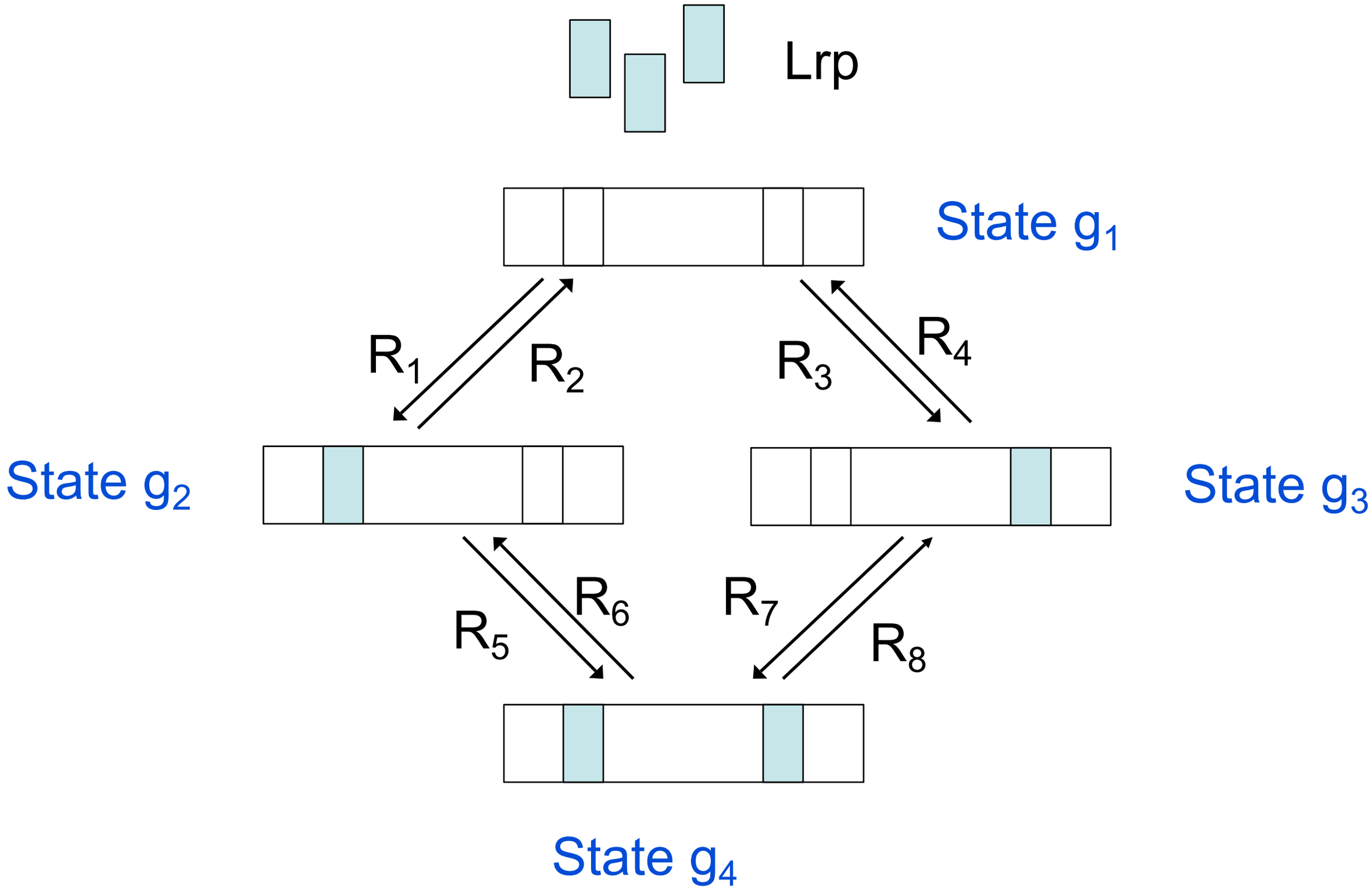


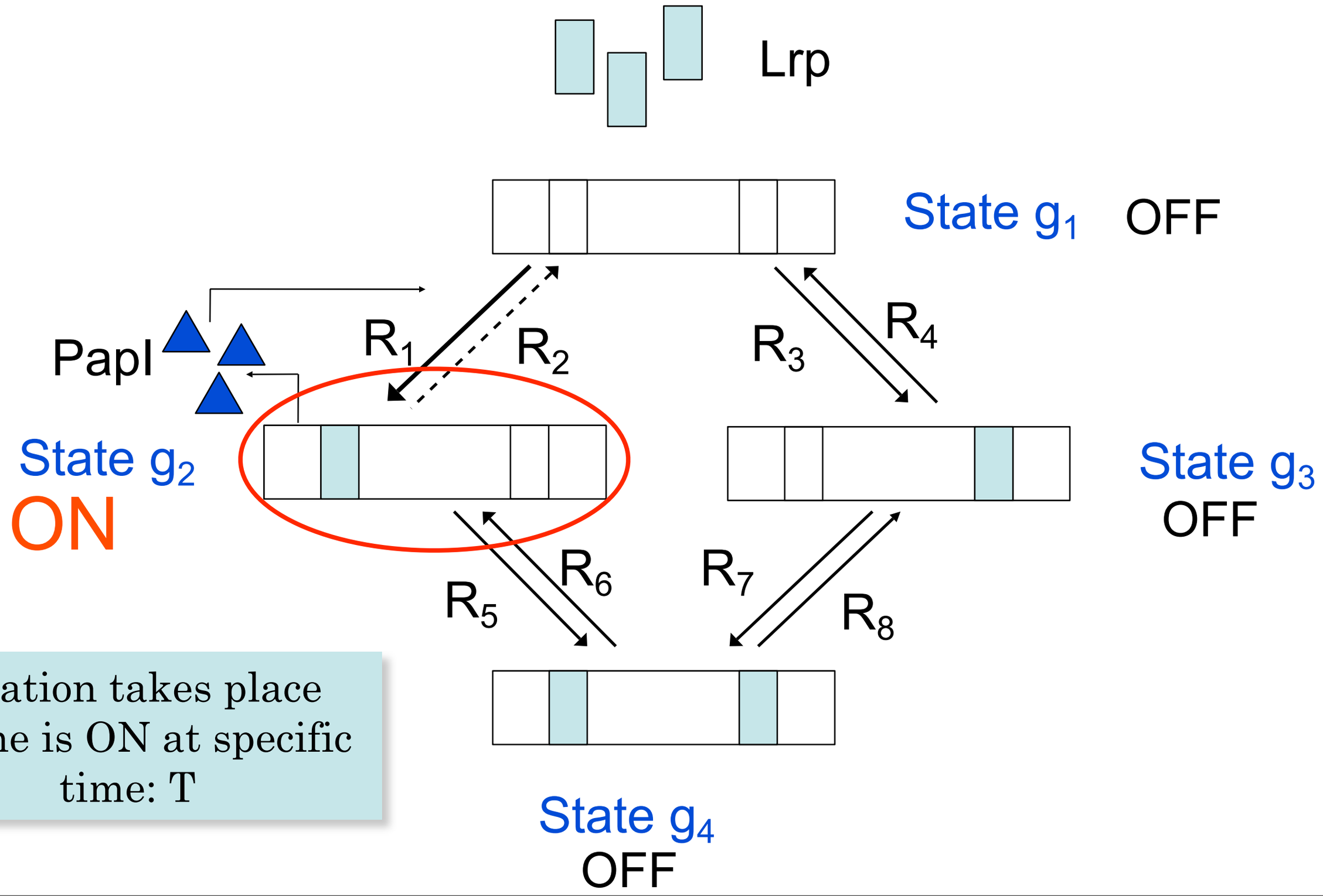
- Pili enable uropathogenic E. coli to attach to epithelial cell receptors
  - Plays an essential role in the pathogenesis of urinary tract infections
- E. coli expresses two states ON (piliated) or OFF (unpiliated)
- Piliation is controlled by a stochastic switch that involves random molecular events

A Simplified Pap Switch Model



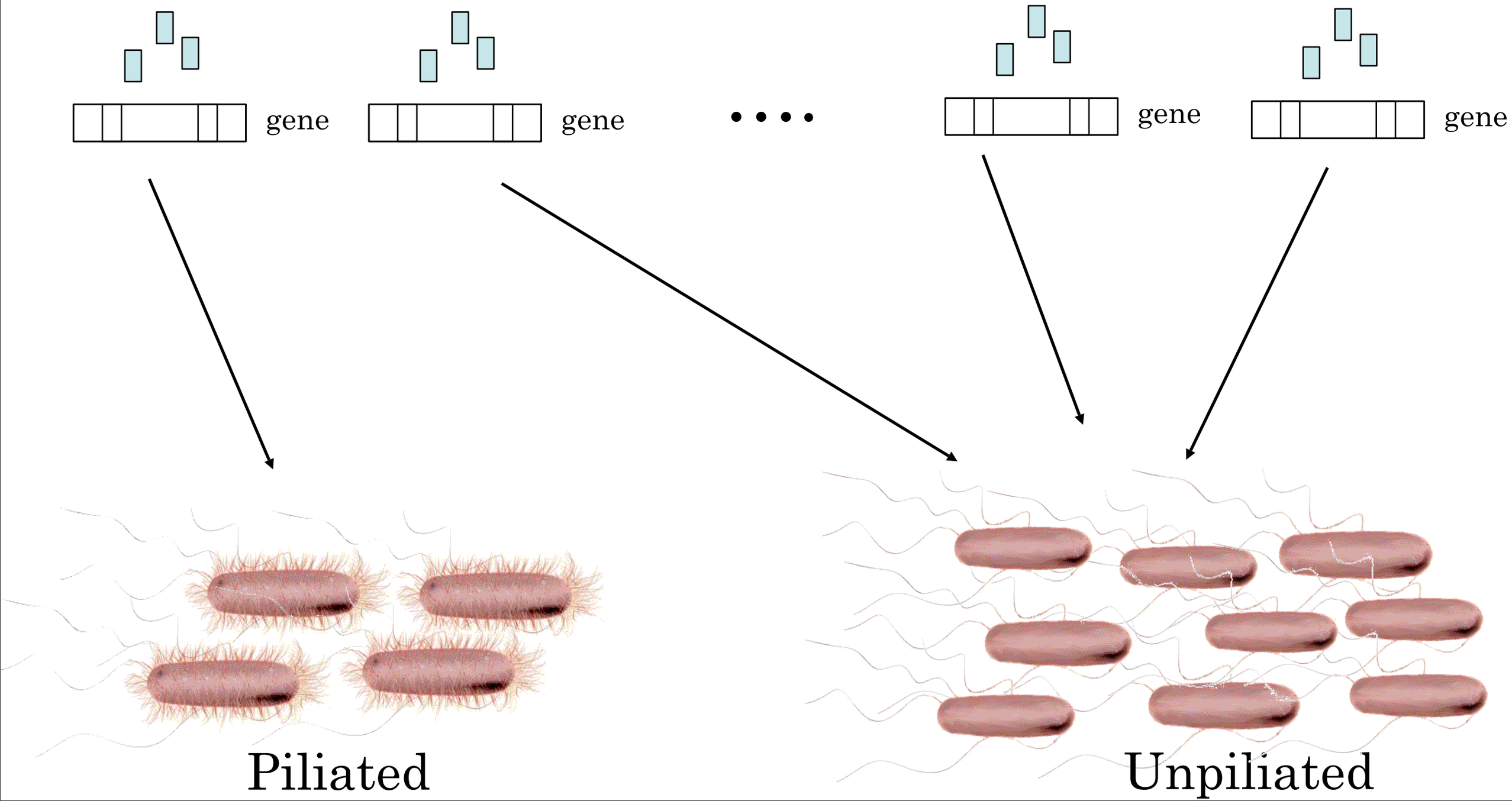
- Lrp can (un)bind either or both of two binding sites
- A (un)binding reaction is a random event


Lrp
State g_1
R_1
R_2
R_3
R_4
State g_2
State g_3
R_5
R_6
R_7
R_8
State g_4



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

Identical Genotype Leads to Different Phenotype



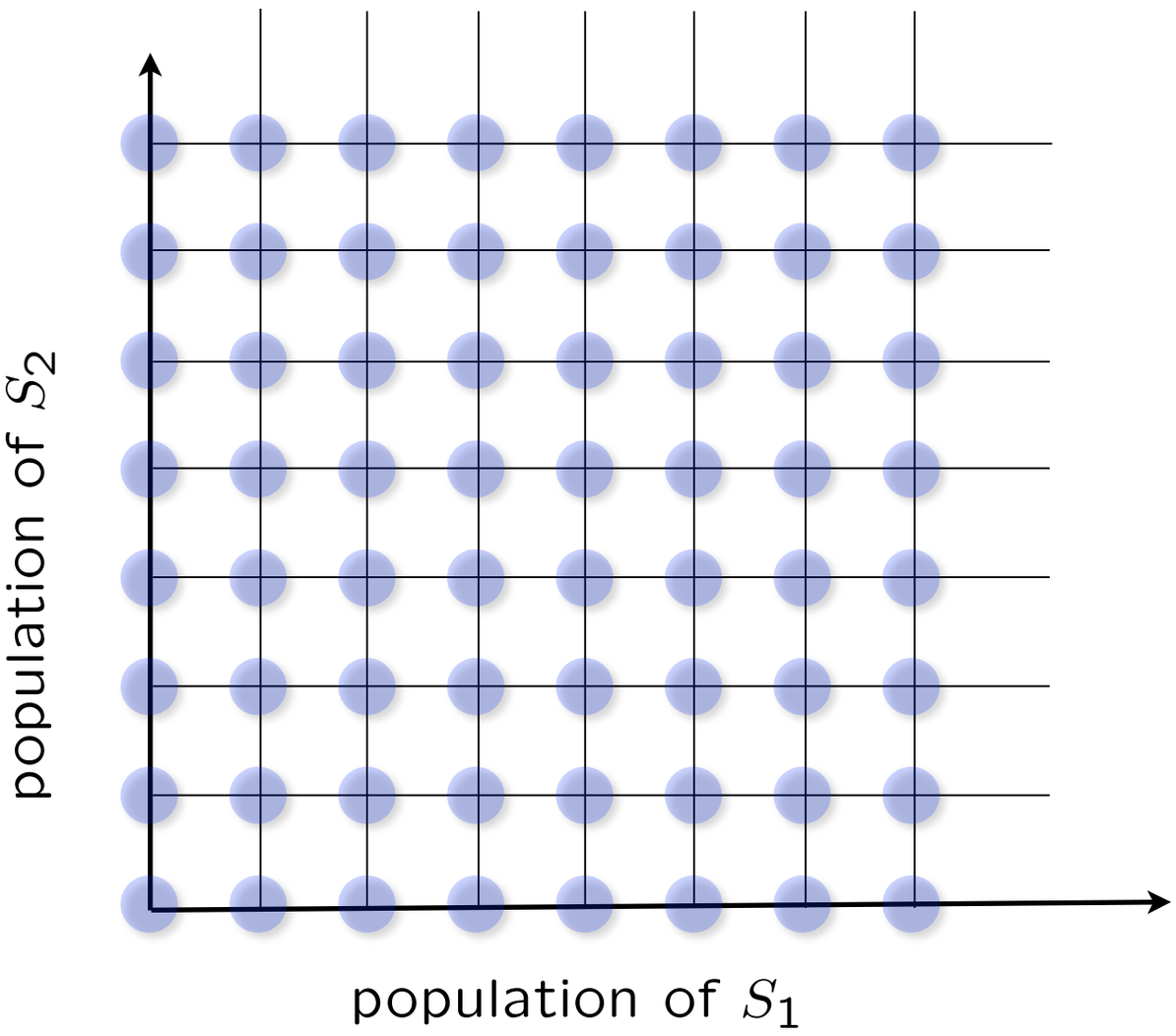
Stochastic Chemical Kinetics

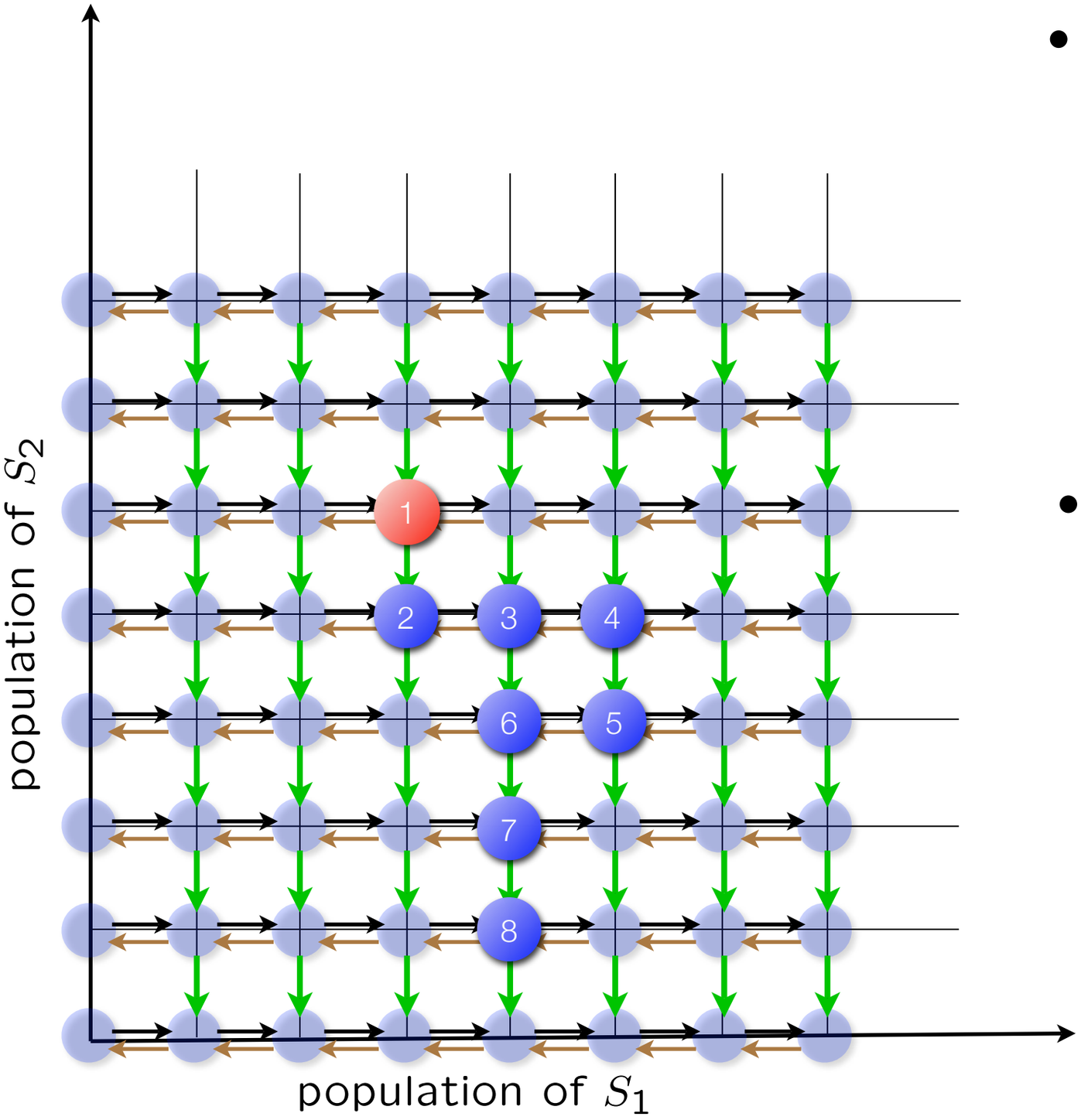
- Mathematical Formulation
- The Chemical Master Equation (CME)
- Monte-Carlo Simulations of the CME

Stochastic Chemical Kinetics

- **(N-species)** Start with a chemically reacting system containing N distinct reacting species $\{S_1, \ldots, S_N\}$.

- The state of the system is described by the integer vector $\mathbf{x} = (x_1, \ldots, x_n)^T$; x_i is the population of species S_i.





- **(M-reactions)** The system's state can change through any one of M reaction: $R_\mu : \mu \in \{1, 2, \ldots, M\}$..

  Example: R_1 $\phi \to S_1$
  R_2 $S_1 + S_2 \to S_1$
  R_3 $S_1 \to \phi$

- **(State transition)** An R_μ reaction causes a state transition from $\mathbf{x}$ to $\mathbf{x} + s_\mu$.

$$s_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad s_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}; \quad s_3 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Stoichiometry matrix:

$$S = \begin{bmatrix} s_1 & s_2 & \cdots & s_M \end{bmatrix}$$

Sequence: R_2 R_1 R_1 R_2 R_3 R_2 R_2

- **(Transition probability)** Suppose the system is in state **x**.

  The probability that R_μ is the next reaction and that it will occurs within the next dt time units is given by $a_\mu(\mathbf{x})dt$.

  If R_μ is the monomolecular reaction $S_i \rightarrow$ products, there exists some constant c_μ such that $a_\mu(\mathbf{x}) = c_\mu x_i$.

  c_μ is numerically equal to the reaction rate constant k_μ of conventional deterministic chemical kinetics.

If R_μ is the bimolecular reaction $S_i + S_j \rightarrow$ products, there exists a constant c_μ such that $a_\mu(\mathbf{x}) = c_\mu x_i x_j$.

c_μ is equal to k_μ/Ω, where k_μ is the reaction rate constant, and Ω is the reaction volume

The Chemical Master Equation

We are interested in $p(\mathbf{x}, t)$, the probability that the chemical system will be in state $\mathbf{x}$ at time, t.

The time evolution of $p(\mathbf{x}, t)$ is described by the **Chemical Master Equation**:

The Chemical Master Equation (CME):

$$\dot{p}(\mathbf{x};t) = -p(\mathbf{x};t)\sum_{\mu=1}^{M} a_\mu(\mathbf{x}) + \sum_{\mu=1}^{M} p(\mathbf{x}-s_\mu;t)a_\mu(\mathbf{x}-s_\mu)$$

Challenges in the Solution of the CME

- The state space is potentially HUGE!

  N species
  Up to n copies/species
  Number of states: n^N

Eukaryotic Cell

$\sim 10^4$ different proteins
$\sim 10^6$ copies/proteins (avg.)
$\sim 10^{6000}$ states!

Proteins bind together to create even more species!

- Very often transitions have vastly different time-scales

- $\|a(\mathbf{x})\| \to \infty$ as $\|\mathbf{x}\| \to \infty$

Exploiting Underlying Biology

- Modularity: Small subsystems can often be isolated
  e.g. heat-shock has ~7 key interacting molecules
- Specificity: Each protein binds to one or very few other molecules

$$N << 10^4$$

N=23 (heat-shock)
N=63 (pap switch)

- Many "key" molecules exist in small numbers: genes, mRNAs, signaling and regulatory proteins, ...
  e.g. liver cells, 97% of mRNAs are present at ~10 copies/cell.
- Key molecules constitute many of the compound species:
  e.g. methylated DNA

For many species: $n << 10^6$

Monte Carlo Simuations: The Stochastic Simulation Algorithm (SSA)

- This approach constructs numerical realizations of $\mathbf{x}(t)$ and then histogramming or averaging the results of many realizations (Gillespie 1976).

- Start by computing the **reaction probability density function** $P_{\mathbf{X}}(\tau, \mu)$:

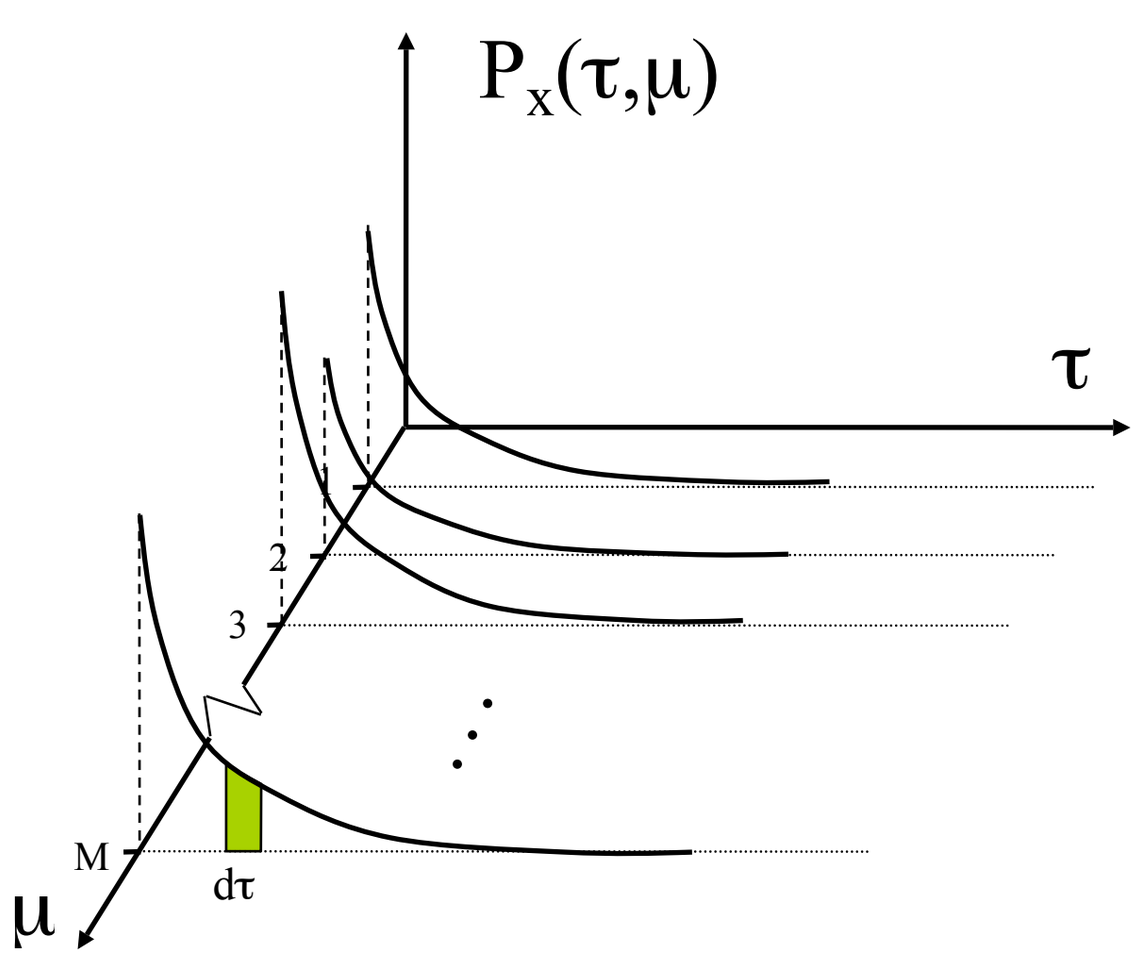
  $P_{\mathbf{X}}(\tau, \mu)$ is the joint probability density function of the two random variables:

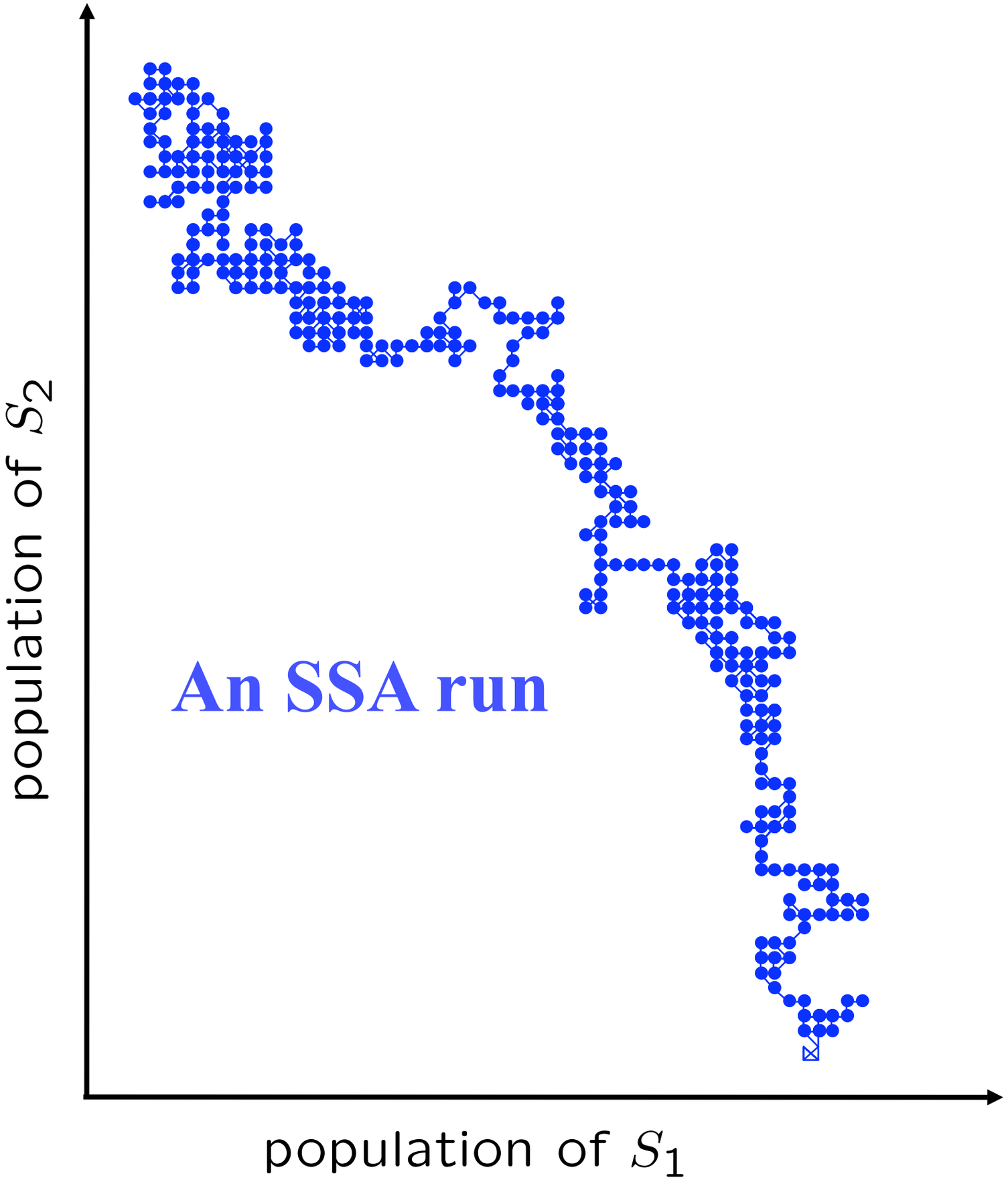
  – "time to the next reaction" (τ);

  – "index of the next reaction" (μ).

$$P_{\mathbf{X}}(\tau, \mu) = a_\mu(\mathbf{x}) \ \exp\left(-\sum_{j=1}^{M} a_j(\mathbf{x})\,\tau\right),$$

- **At each SSA step:** Based on this distribution, the next reaction and the time of its occurence are determined. The state is updated.




population of S_2
An SSA run
population of S_1

SSA

Advantages

- Low memory requirement
- In principle, computation does not depend exponentially on N

Disadvantages

- Can be very slow
- Convergence is slow $\quad \text{Computational effort} = C\frac{1}{\epsilon^2}$
- No guaranteed error bounds
- Little insight

A Direct Approach to the Chemical Master Equation

- The Finite State Projection Method
- Model Reduction and Aggregation
- Examples

We are interested in $p(\mathbf{x}, t)$, the probability that the chemical system will be in state $\mathbf{x}$ at time, t.

The Chemical Master Equation (CME):

$$\dot{p}(\mathbf{x};t) = -p(\mathbf{x};t)\sum_{\mu=1}^{M} a_\mu(\mathbf{x}) + \sum_{\mu=1}^{M} p(\mathbf{x}-\nu_\mu;t)a_\mu(\mathbf{x}-\nu_\mu)$$

The states of the chemical system can be enumerated:

$$\mathbf{X} := [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \mathbf{x}_3 \quad \dots \quad]^T$$

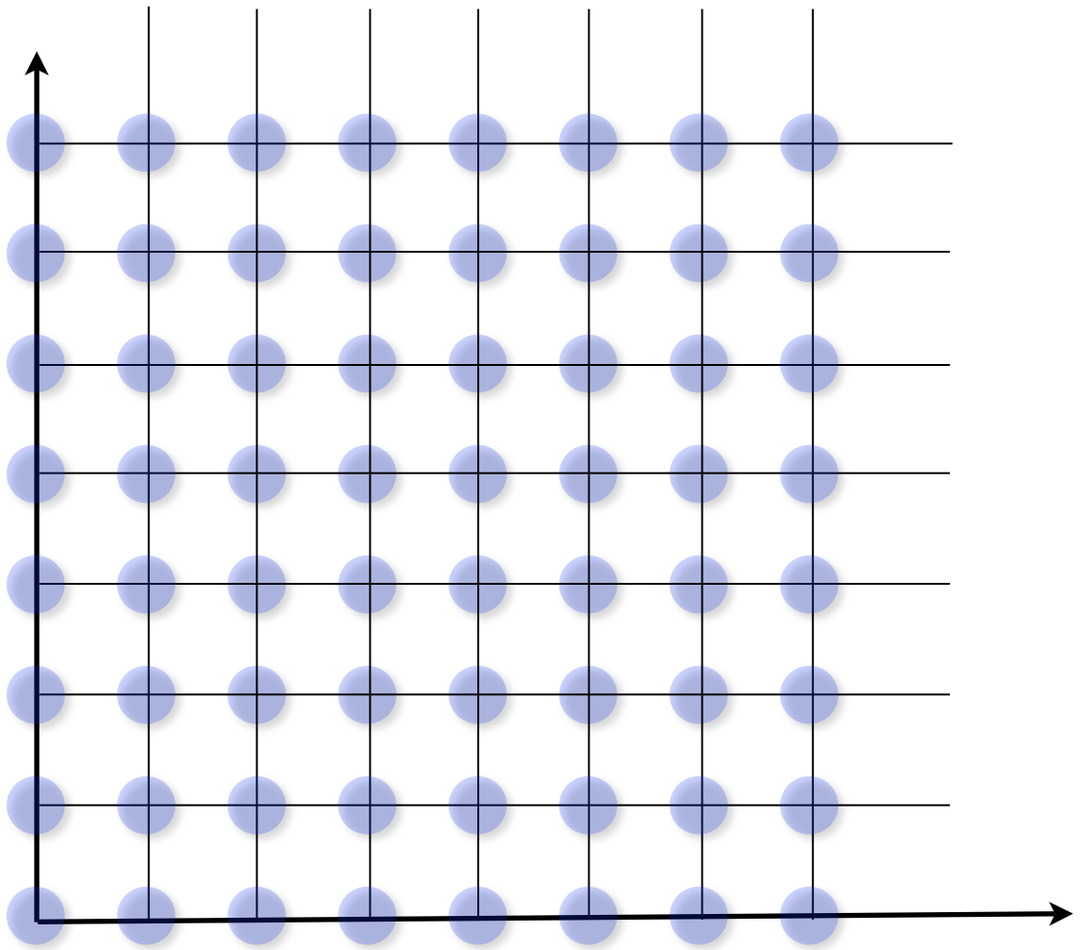
The *probability density state vector* $\mathbf{P}(\mathbf{X}, \cdot) : R \to \ell_1$ defined by:

$$\mathbf{P}(\mathbf{X};t) := [p(\mathbf{x}_1;t) \quad p(\mathbf{x}_2;t) \quad p(\mathbf{x}_3;t) \quad \dots \quad]^T$$

The evolution of the *probability density* state vector is governed by

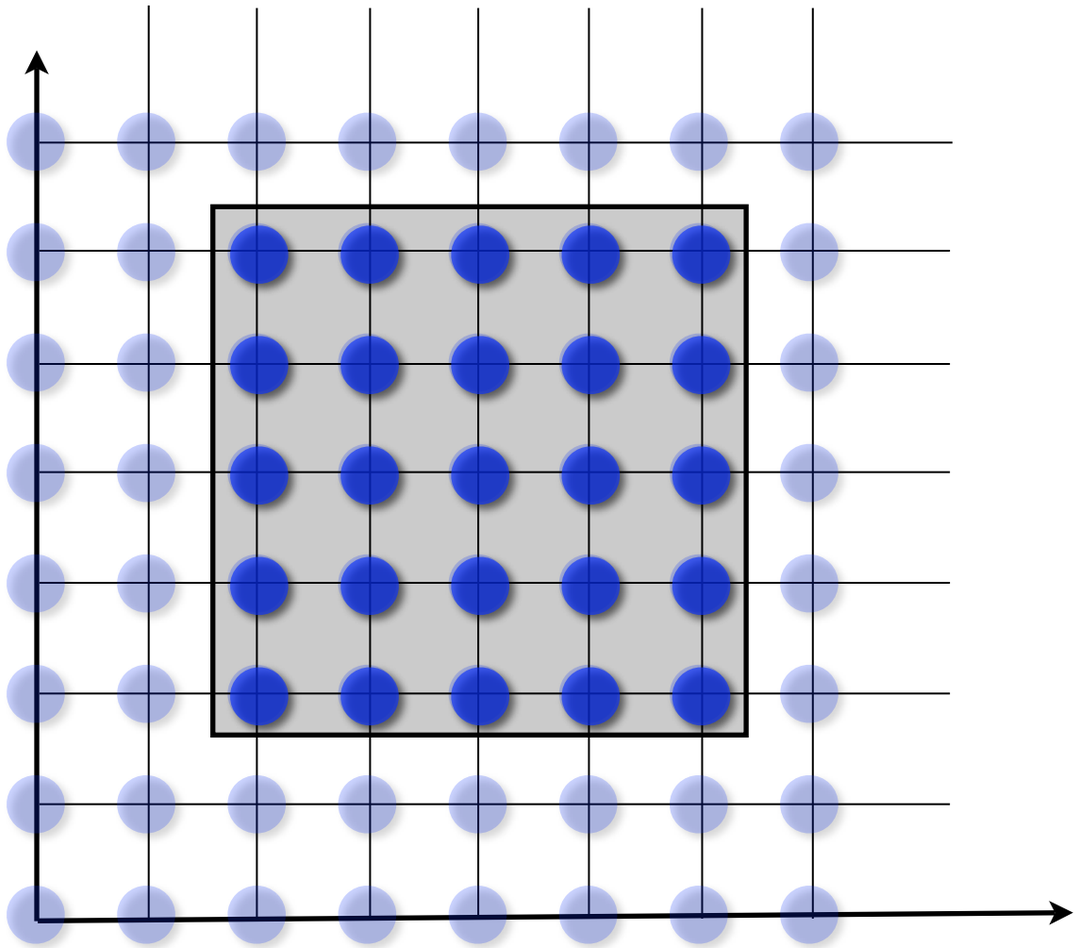
$$\dot{\mathbf{P}}(\mathbf{X};t) = \mathbf{A} \cdot \mathbf{P}(\mathbf{X};t)$$

The Finite State Projection Approach



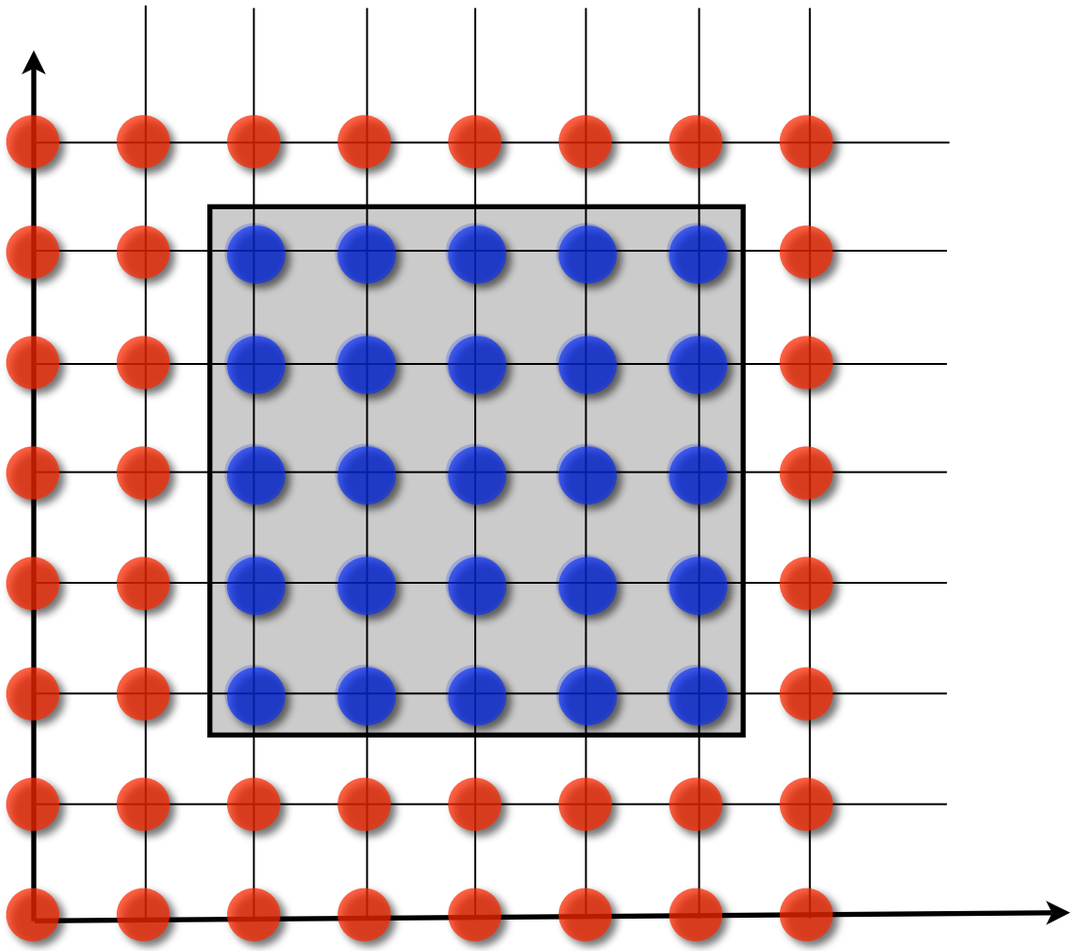
The probability density vector evolves on an infinite integer lattice

The Finite State Projection Approach



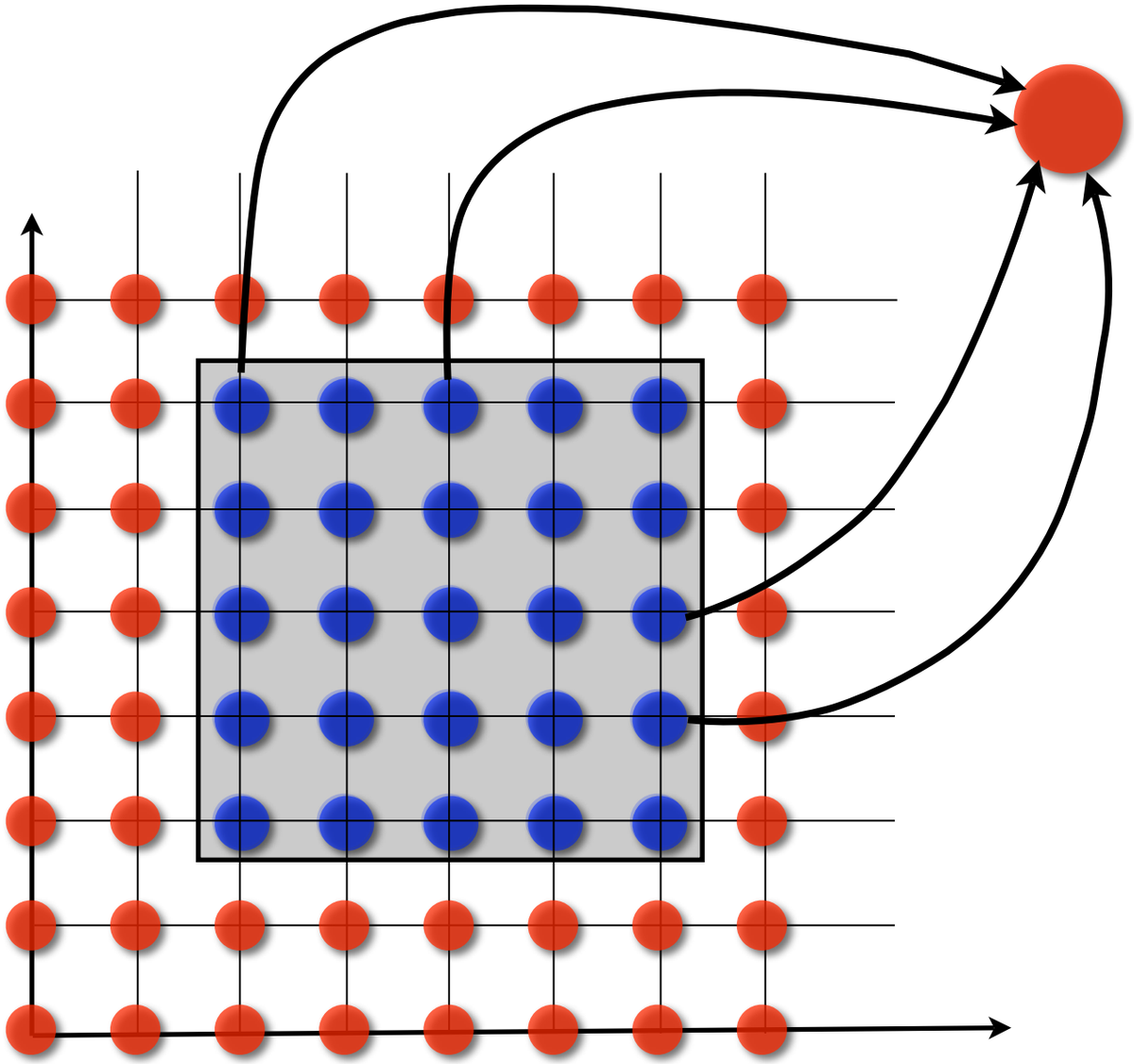
- A finite subset is appropriately chosen

The Finite State Projection Approach



- A finite subset is appropriately chosen
- The remaining (infinite) states are projected onto a single state (red)

The Finite State Projection Approach



- A finite subset is appropriately chosen
- The remaining (infinite) states are projected onto a single state (red)
- Only transitions into removed states are retained

The projected system can be solved exactly!

Finite Projection Bounds

Let $J = [m_1 \ldots m_N]$ be an indexing vector. We define $\mathbf{A}_J$ to be the principle submatrix of $\mathbf{A}$ defined by J.

Theorem [Projection Error Bound]: Consider any Markov process in which the probability distribution evolves according to the ODE:

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

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

$$\left\| \begin{bmatrix} \mathbf{P}(\mathbf{X}_J; t) \\ \mathbf{P}(\mathbf{X}_{J'}; t) \end{bmatrix} - \begin{bmatrix} \exp(\mathbf{A}_J t)\ \mathbf{P}(\mathbf{X}_J; 0) \\ 0 \end{bmatrix} \right\|_1 \leq \epsilon$$

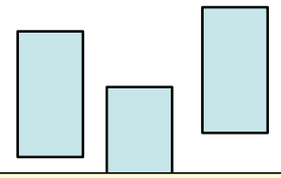
Munsky B. and Khammash M., Journal of Chemical Physics, 2006

The FSP Algorithm

- **Step 0.** Define the propensity functions and stoichiometry for all reactions.
  - Choose the initial probability density function $\mathbf{P}(\mathbf{X}, 0)$.
  - Choose the final time of interest, t.
  - Choose the total amount of acceptable error ϵ.
  - Choose an initial finite set of states: $\mathbf{X}_{J_0}$.
  - Set $i = 0$.
- **Step 1.** Form $\mathbf{A}_{J_i}$. Compute $\Gamma_{J_i} = \mathbf{1}^T \exp(\mathbf{A}_{J_i} t)\ \mathbf{P}(\mathbf{X}_{J_i}; 0)$.
- **Step 2.** If $\Gamma_{J_i} \geq 1 - \epsilon$: stop.

  $\exp(\mathbf{A}_{J_i} t)\ \mathbf{P}(\mathbf{X}_{J_i}; 0)$ approximates $\mathbf{P}(\mathbf{X}_{J_i}; t)$ to within ϵ.
- **Step 3.** Add more states to get $\mathbf{X}_{J_{i+1}}$. Increment i. Go to step 1.

Application to the Pap Switch Model



Lrp

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

State g_2 ON

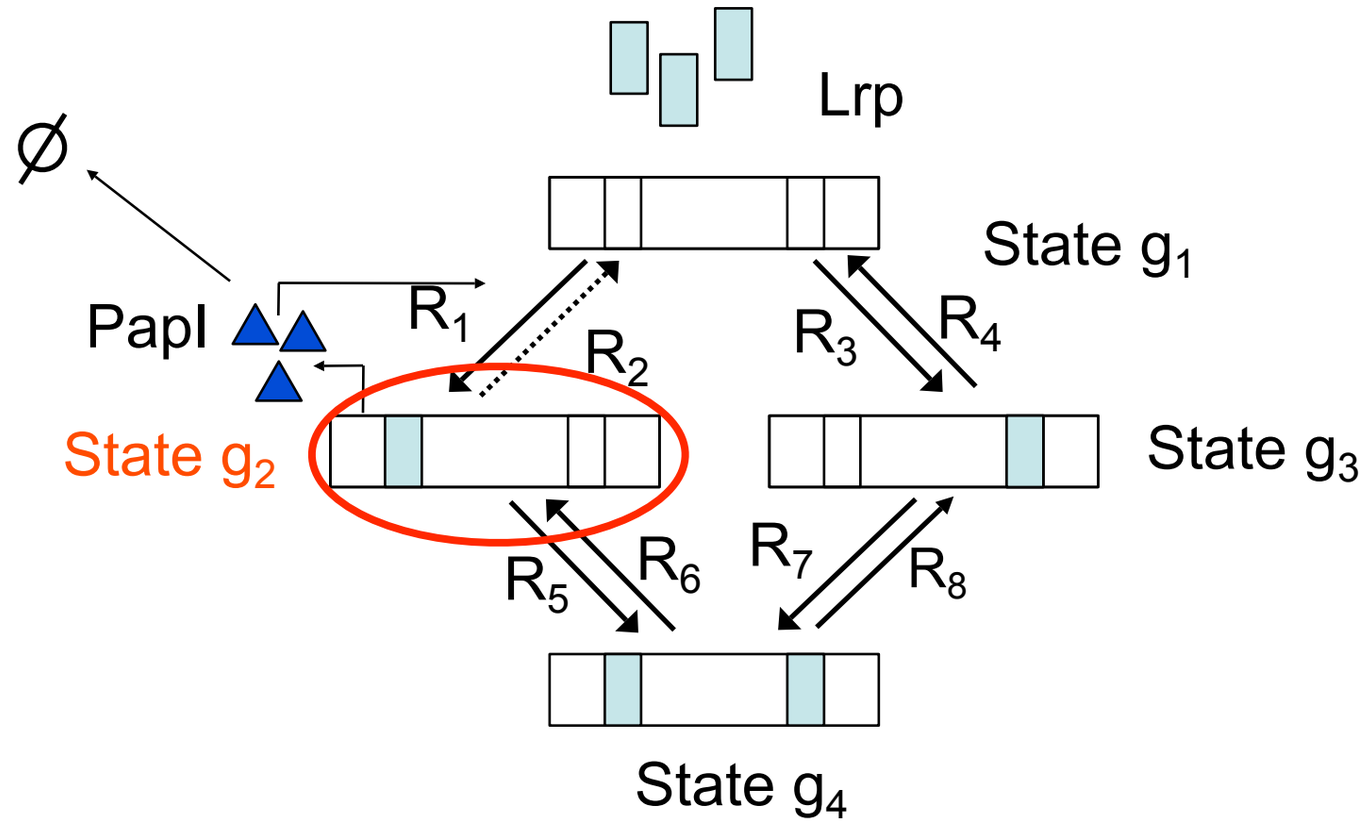
State g_3 OFF

R_5 R_6 R_7 R_8

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

State g_4 OFF

Application to the Pap Switch Model



10 reactions

An *infinite* number of states:

$$\left\{ \begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \\ LRP \\ PapI \end{bmatrix}_i \right\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ u_0 \\ i \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ u_0 - 1 \\ i \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ u_0 - 1 \\ i \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ u_0 - 2 \\ i \end{bmatrix} \right\}$$

$$i = 0, 1, \ldots$$

Application to the Pap Switch Model

Reactions			
Number	Stoichiometry	Rate Constant(c_μ)	Units
R_1	$\mathbf{X}_1 + u \rightarrow \mathbf{X}_2$	1	s^{-1}
R_2	$\mathbf{X}_2 \rightarrow \mathbf{X}_1 + u$	$2.5 - 2.25\left(\frac{r}{1+r}\right)$	s^{-1}
R_3	$\mathbf{X}_1 + u \rightarrow \mathbf{X}_3$	1	s^{-1}
R_4	$\mathbf{X}_3 + u \rightarrow \mathbf{X}_1$	$1.2 - 0.2\left(\frac{r}{1+r}\right)$	s^{-1}
R_5	$\mathbf{X}_2 + u \rightarrow \mathbf{X}_4$	0.01	s^{-1}
R_6	$\mathbf{X}_4 + u \rightarrow \mathbf{X}_2$	$1.2 - 0.2\left(\frac{r}{1+r}\right)$	s^{-1}
R_7	$\mathbf{X}_3 + u \rightarrow \mathbf{X}_4$	0.01	s^{-1}
R_8	$\mathbf{X}_4 + u \rightarrow \mathbf{X}_3$	$2.5 - 2.25\left(\frac{r}{1+r}\right)$	s^{-1}
R_T	$\mathbf{X}_2 \rightarrow \mathbf{X}_2 + r$	10	s^{-1}
R_D	$r \rightarrow \emptyset$	r	s^{-1}

Parameters and Initial Conditions		
Parameter	Notation	Value
Initial Catalyst Protein	r	5
Initial Gene State	g_1	–
Initial *pdf*	$\mathbf{P}(\mathbf{X};0)_{21} = 1$	–
Allowable Error in *pdf*	ϵ	10^{-5}

A. Hernday and B. Braaten, and D. Low, “The Mechanism by which DNA Adenine Methylase and PapI Activate the Pap Epigenetic Switch,” Molecular Cell,vol. 12, 947-957, 2003.

Application to the Pap Switch Model

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
τ leaping	1.25×10^5	≈18	9.9
τ leaping	2.5×10^5	≈35	24.4
τ leaping	5.0×10^5	≈70	7.0
τ leaping	10.0×10^5	≈140	6.0

- Unlike Monte-Carlo methods (such as the SSA), the FSP *directly* approximates the solution of the CME.

Model Reduction and Time-Scale Separation

Reducing Unobservable Configurations

- Often one is not interested in the entire probability distribution. Instead one may wish only to estimate:
  - a statistical summary of the distribution, e.g.
    - means, variances, or higher moments
  - probability of certain traits:
    - switch rate, extinction, specific trajectories, etc…
- In each of these cases, one can define an output y:

$$\dot{\mathbf{P}}(t) = \mathbf{A}\mathbf{P}(t)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{P}(t)$$

- Frequently, the output of interest relates to small subset of the state space

Aggregation and Model Reduction

Given Generic CME in the form of a linear ODE:

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

The system begins in the set U at time $t = 0$ with pdv: $\mathbf{P}(\mathbf{X}_U, 0)$.

Find: $\mathbf{P}(\mathbf{X}_Y, t_f)$, for some set Y.

Define:

$R =$ set of states reachable from the U.

$R' =$ set of states unreachable from U.

$O =$ set of states from which Y may be reached.

$O' =$ set of states unobservable from Y.

Aggregation and Model Reduction

The full pdv evolves according to:

$$\begin{bmatrix} \dot{\mathbf{P}}(\mathbf{X}_{RO},t) \\ \dot{\mathbf{P}}(\mathbf{X}_{R'O},t) \\ \dot{\mathbf{P}}(\mathbf{X}_{RO'},t) \\ \dot{\mathbf{P}}(\mathbf{X}_{R'O'},t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{RO,RO} & \mathbf{A}_{RO,R'O} & 0 & 0 \\ 0 & \mathbf{A}_{R'O,R'O} & 0 & 0 \\ \mathbf{A}_{RO',RO} & \mathbf{A}_{RO',R'O} & \mathbf{A}_{RO',RO'} & \mathbf{A}_{RO',R'O'} \\ 0 & \mathbf{A}_{R'O',R'O} & 0 & \mathbf{A}_{R'O',R'O'} \end{bmatrix} \begin{bmatrix} \mathbf{P}(\mathbf{X}_{RO},t) \\ \mathbf{P}(\mathbf{X}_{R'O},t) \\ \mathbf{P}(\mathbf{X}_{RO'},t) \\ \mathbf{P}(\mathbf{X}_{R'O'},t) \end{bmatrix}$$

- The unreachable states cannot be excited by reachable ones (may be removed!)
- The unobservable states may not excite the observable ones

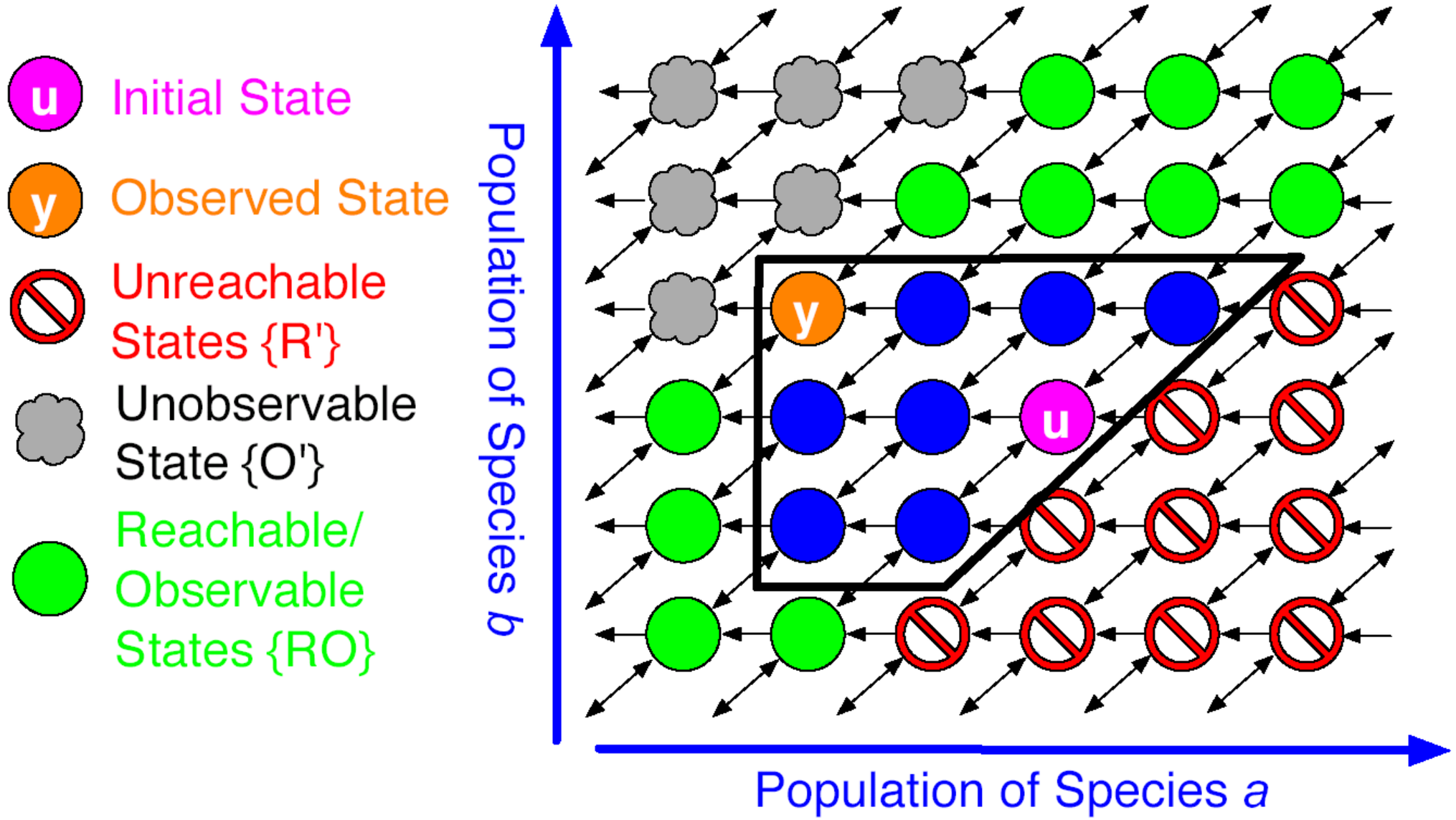
The full system reduces to:

$$\begin{bmatrix} \dot{\mathbf{P}}(\mathbf{X}_{RO},t) \\ \dot{\mathbf{P}}(\mathbf{X}_{RO'},t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{RO,RO} & \mathbf{0} \\ \mathbf{A}_{RO',RO} & \mathbf{A}_{RO',RO'} \end{bmatrix} \begin{bmatrix} \mathbf{P}(\mathbf{X}_{RO},t) \\ \mathbf{P}(\mathbf{X}_{RO'},t) \end{bmatrix}$$

We need only keep track of the dynamics of unobservable states as a whole

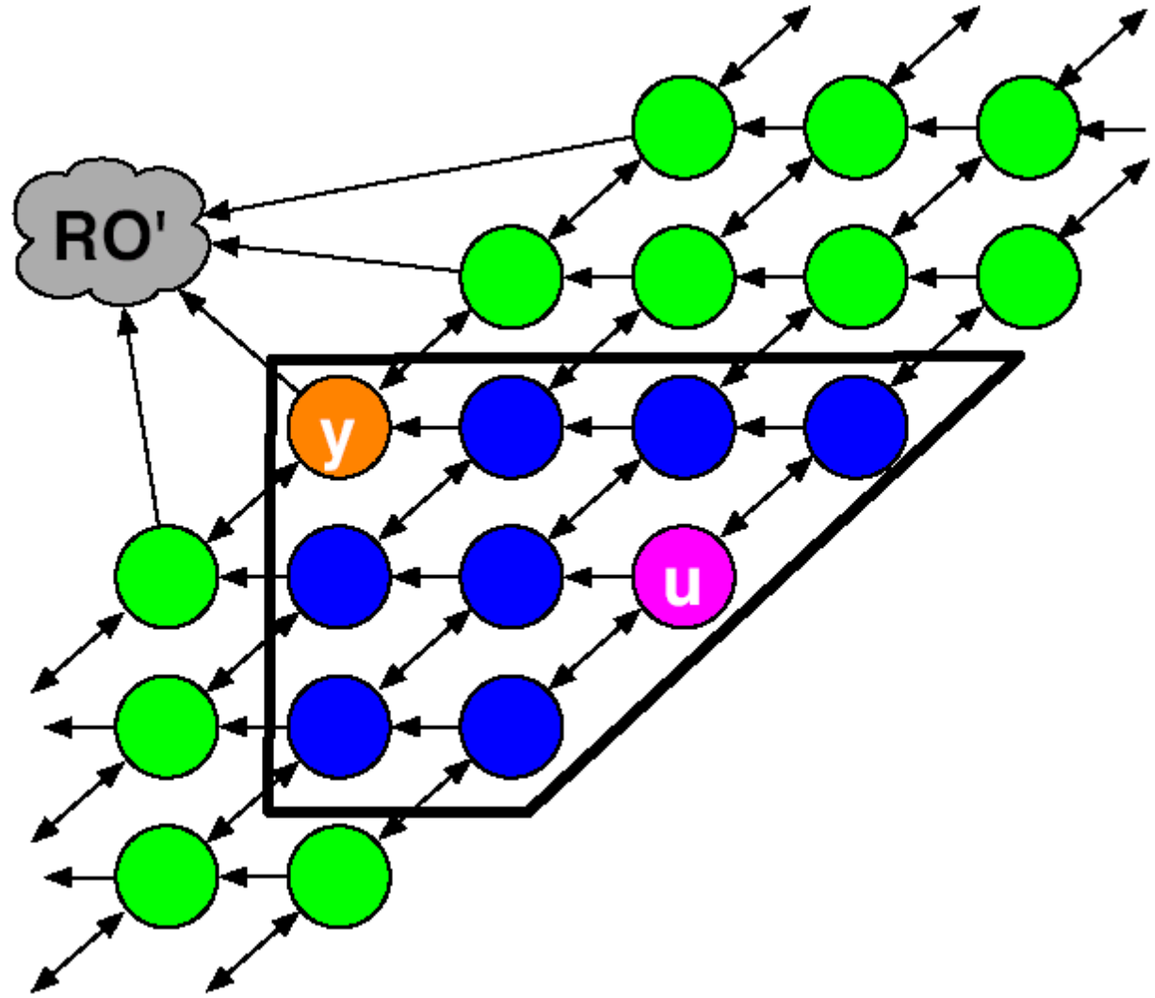
$$\begin{bmatrix} \dot{\mathbf{P}}(\mathbf{X}_{RO},t) \\ \mathbf{1}^T\dot{\mathbf{P}}(\mathbf{X}_{RO'},t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{RO,RO} & \mathbf{0} \\ \mathbf{1}^T\mathbf{A}_{RO',RO} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{P}(\mathbf{X}_{RO},t) \\ \mathbf{P}(\mathbf{X}_{RO'},t) \end{bmatrix}$$

FSP: Aggregation and Model Reduction



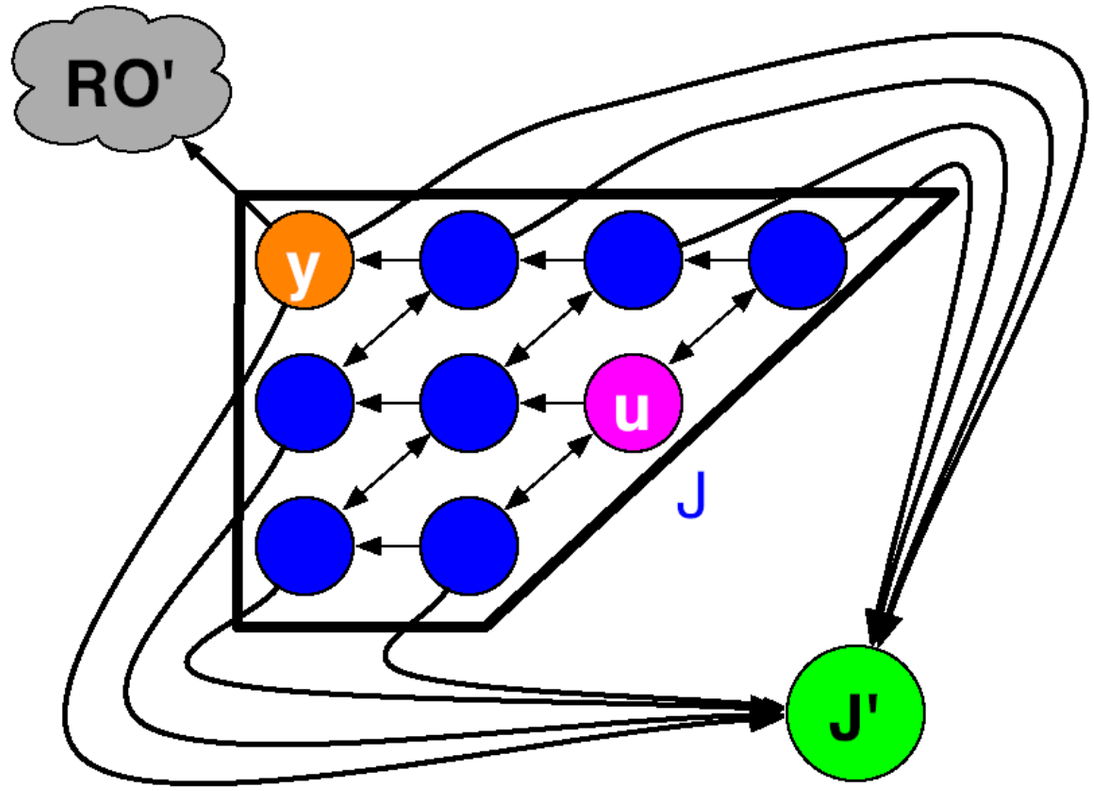
Begin with a full integer lattice description of the system

FSP: Aggregation and Model Reduction



Remove unreachable states and aggregate unobservable ones

FSP: Aggregation and Model Reduction



Project the remaining states onto a finite subset

Reduction Through Time-Scale Separation

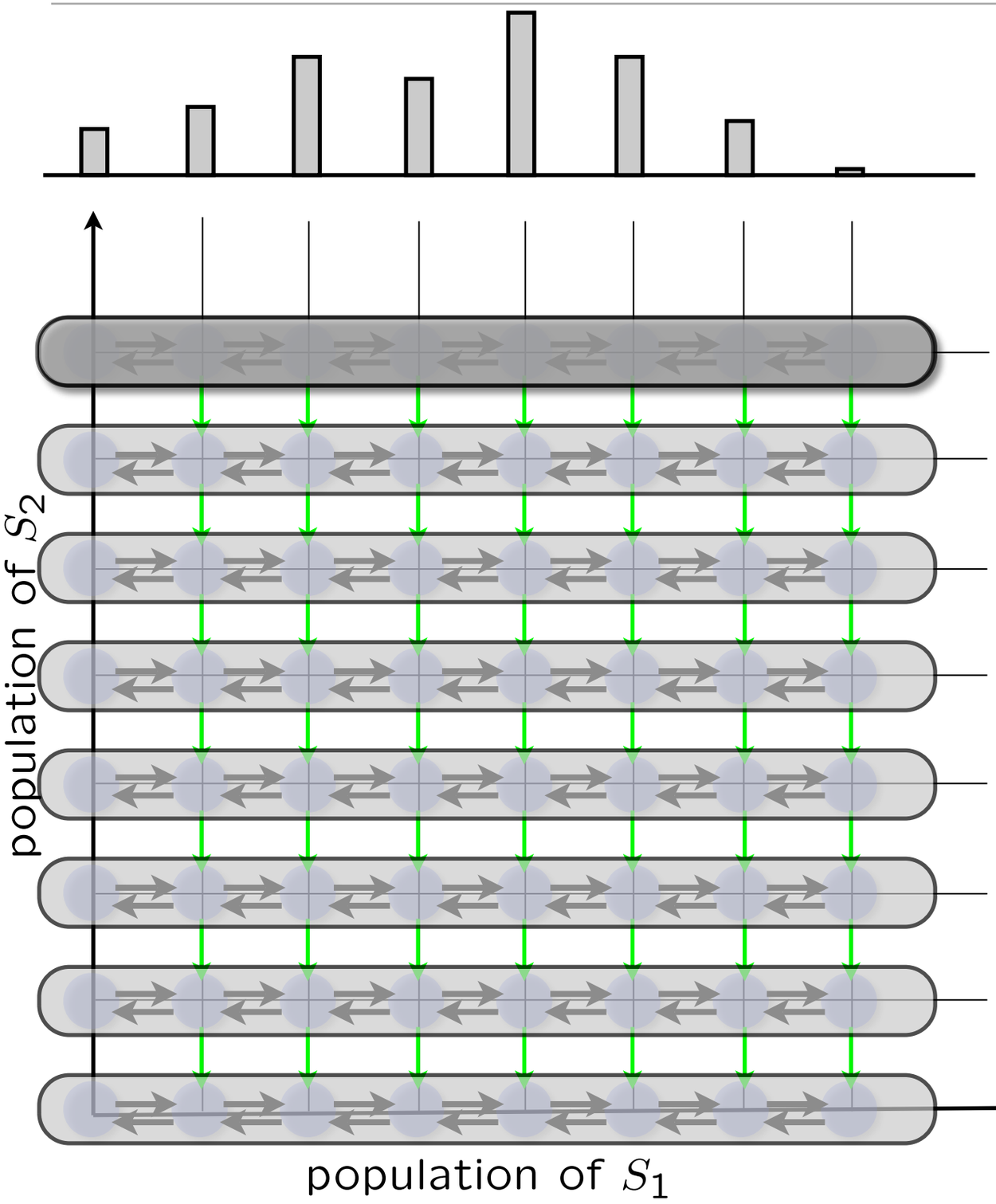
- It often occurs in chemical systems that some reactions are very fast, while others are relatively slow

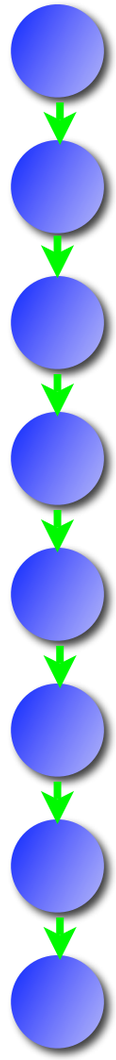
- These differences in reaction rates lead to significant numerical stiffness

- Often the fast dynamics are not of interest

- One may remove the fast dynamics and focus on the dynamics on the slow manifold

Time Scale Separation





Group together states that are connected through fast transitions

Fast groups reach probabilistic equilibrium before a slow transition occurs

Aggregate fast group into states

Transition propensity is the weighted sum of transition propensities of unaggregated states

Time Scale Separation

$$\dot{\mathbf{P}} = \left(\begin{bmatrix} \mathbf{H}_1 & \mathbf{0} & & \\ \mathbf{0} & \mathbf{H}_2 & \ddots & \\ & \ddots & \ddots & \\ & & \ddots & \mathbf{H}_m \end{bmatrix} + \varepsilon \mathbf{L} \right) \mathbf{P}$$

- Each H_i is a generator for a fast group of states. ϵL is the slow coupling reactions between blocks.

- Each H_i has a single zero eigenvalue, $\lambda = 0$, the rest of the eigenvalues have large negative real parts.

- Each H_i has a right eigenvector, v_i , and a left eigenvector, $\mathbf{1}^T$ that correspond to the zero eigenvalue.

Time Scale Separation

- Collect the right and left zero eigenvectors

$$\mathbf{U} = \begin{bmatrix} \mathbf{1}^T & \mathbf{0} & \\ \mathbf{0} & \mathbf{1}^T & \ddots \\ & \ddots & \ddots \\ & & \ddots\, \mathbf{1}^T \end{bmatrix} \qquad \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{0} & \\ \mathbf{0} & \mathbf{v}_2 & \ddots \\ & \ddots & \ddots \\ & & \ddots\, \mathbf{v}_m \end{bmatrix}$$

- Define the similarity transformation:

$$\mathbf{S} = \begin{bmatrix} \mathbf{U} \\ \mathbf{S}_1 \end{bmatrix} \qquad \mathbf{S}^{-1} = \begin{bmatrix} \mathbf{V} & \mathbf{S}_2 \end{bmatrix} \qquad \mathbf{SHS}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{bmatrix}$$

- In the new basis, the system becomes:

$$\begin{bmatrix} \dot{\mathbf{q}}_1 \\ \dot{\mathbf{q}}_2 \end{bmatrix} = \begin{bmatrix} \varepsilon\mathbf{ULV} & \varepsilon\mathbf{ULS}_2 \\ \varepsilon\mathbf{S}_1\mathbf{LV} & \mathbf{\Lambda}_2 + \varepsilon\mathbf{S}_1\mathbf{LS}_2 \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{bmatrix}$$

Time Scale Separation

$$\begin{bmatrix} \dot{\mathbf{q}}_1 \\ \dot{\mathbf{q}}_2 \end{bmatrix} = \begin{bmatrix} \varepsilon\mathbf{ULV} & \varepsilon\mathbf{ULS}_2 \\ \varepsilon\mathbf{S}_1\mathbf{LV} & \mathbf{\Lambda}_2 + \varepsilon\mathbf{S}_1\mathbf{LS}_2 \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{bmatrix}$$

- Ignoring the fast stable modes, the dynamics of q_1

$$\dot{\mathbf{q}}_1 \approx \varepsilon\mathbf{ULV}\mathbf{q}_1$$

- In the original coordinate system, $P(t)$ is given by

$$\mathbf{P}(t) \approx \varepsilon\mathbf{V}\,\mathrm{exp}(\mathbf{ULV}t)\mathbf{UP}(0)$$

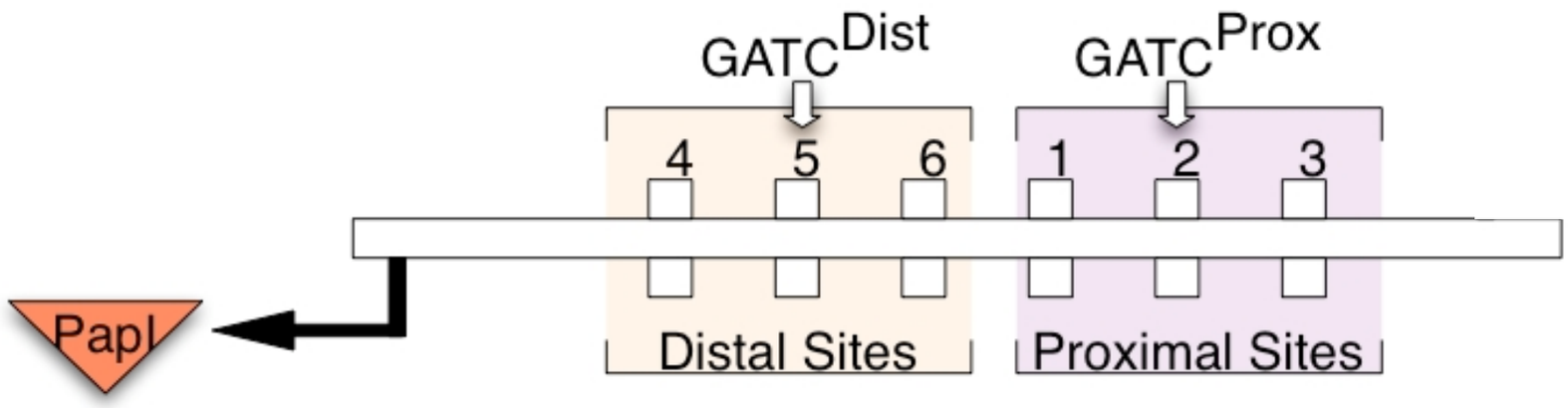
- Note that only the zero eigenvectors of H_i need to be computed!

Simon HA, Ando A. Aggregation of variables in dynamic systems. Econometrica 1961; 29(2):111–138.

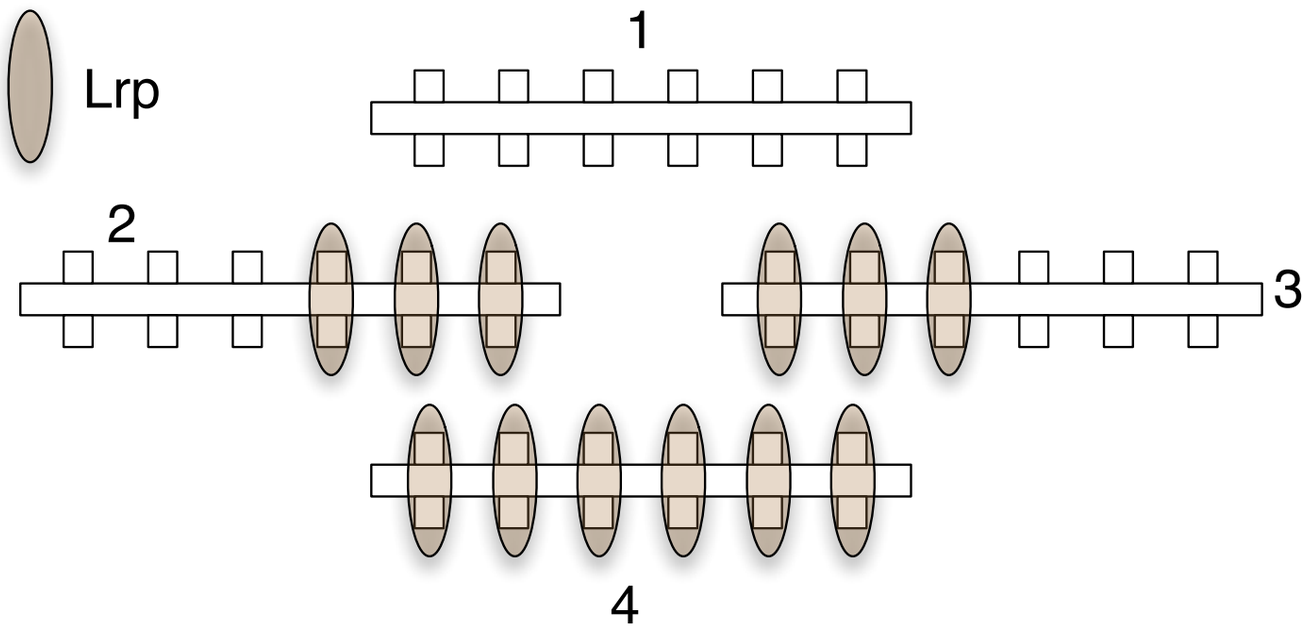
Phillips RG, and Kokotovic P, A Singular perturbation approach to modeling and control of Markov Chains, *IEEE Transactions on Automatic Control*, 26 (5): 1087-1094, 1981.

Example: The Full Pap Switch Model

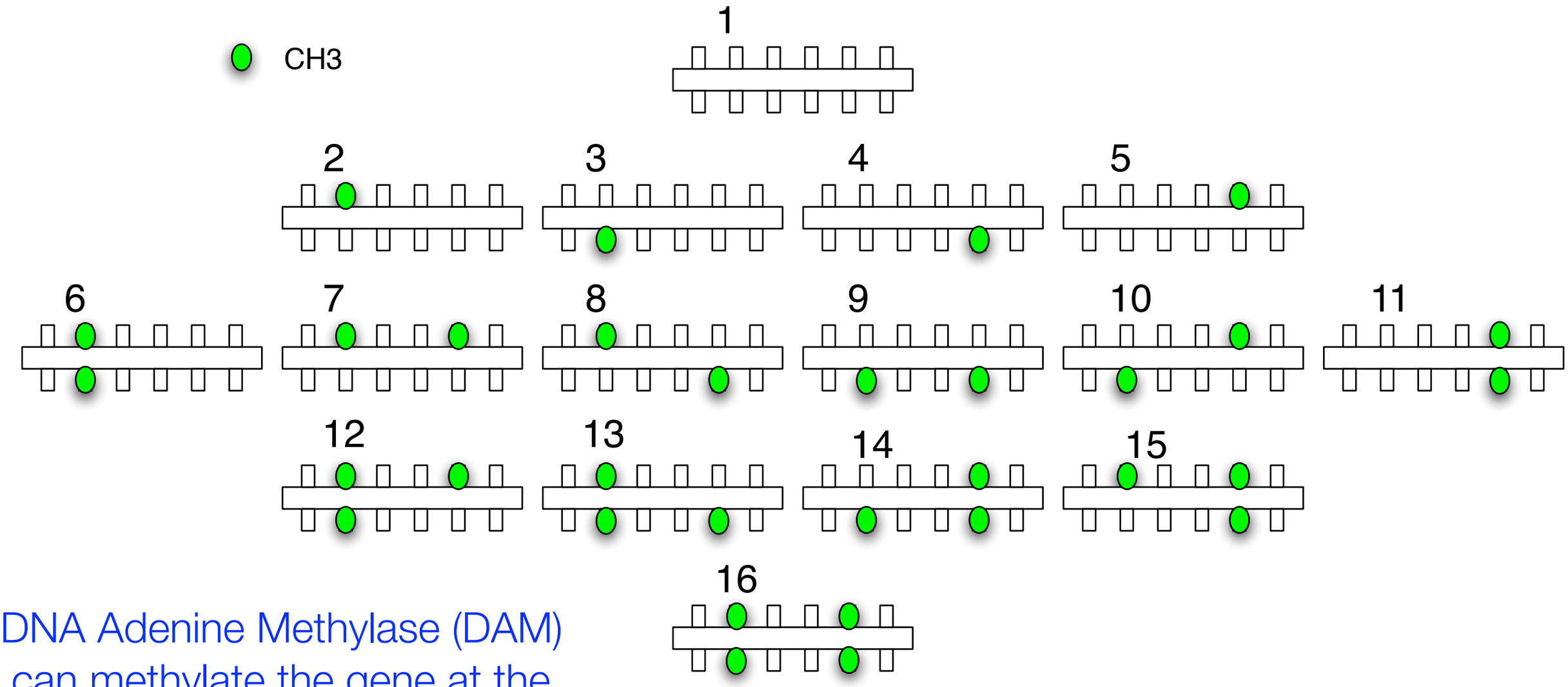
A More Realistic Pap Switch Model



4 gene states based on Lrp binding sites



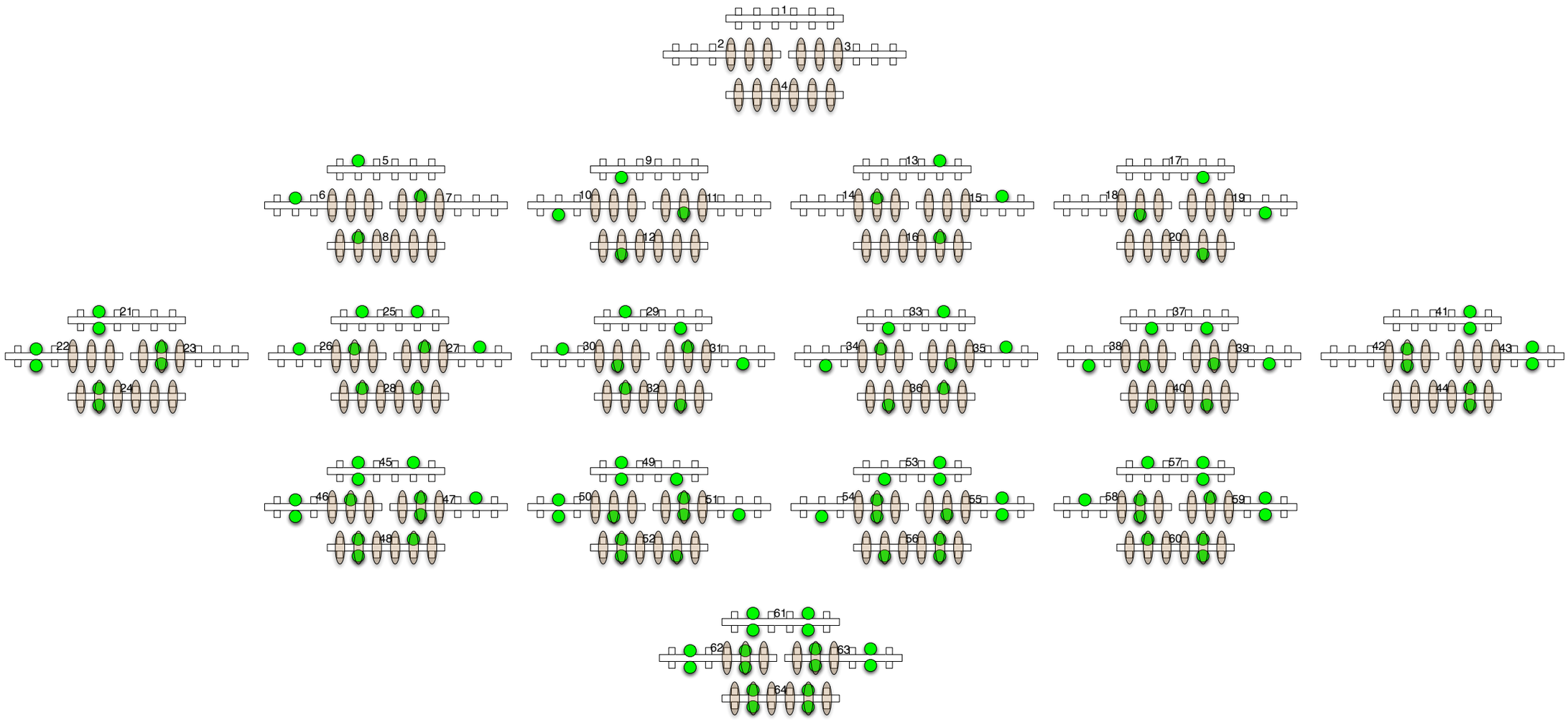
A More Realistic Pap Switch Model



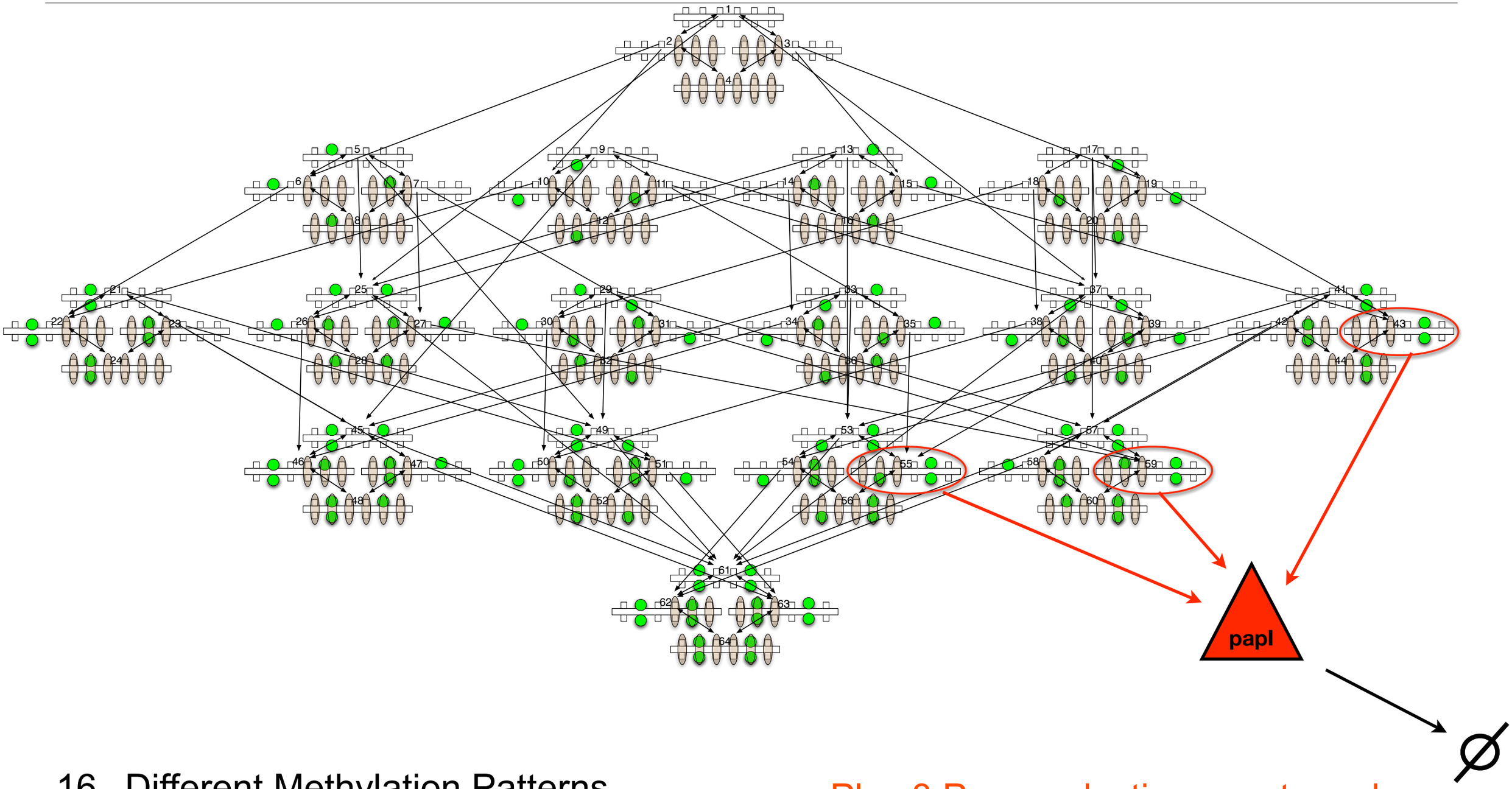
DNA Adenine Methylase (DAM) can methylate the gene at the GATC site

16 different possible methylation patterns

A More Realistic Pap Switch Model



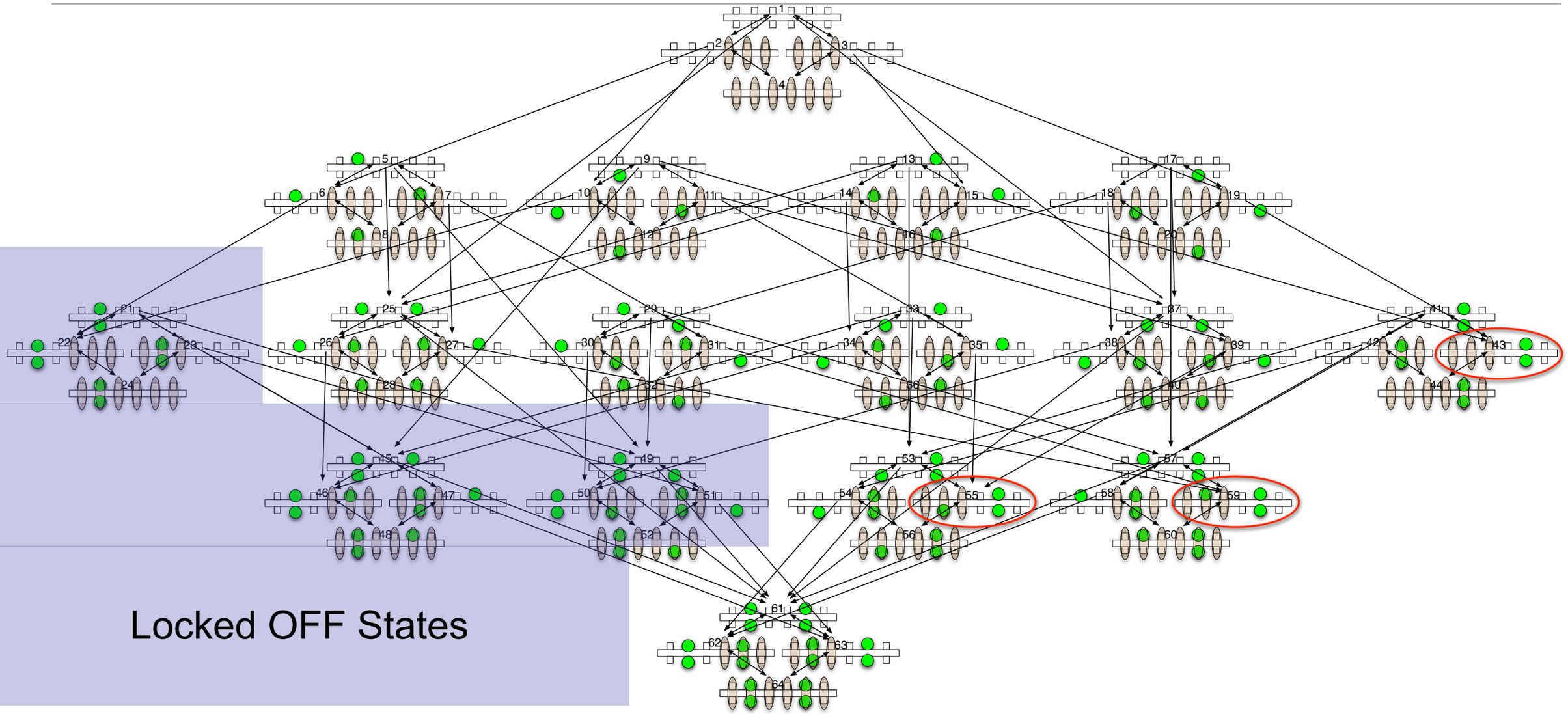
A More Realistic Pap Switch Model



16 Different Methylation Patterns
x 4 Different LRP binding Patterns
=64 Different Operon Configurations!

Plus 3 Pap production events and one Pap degradation event.

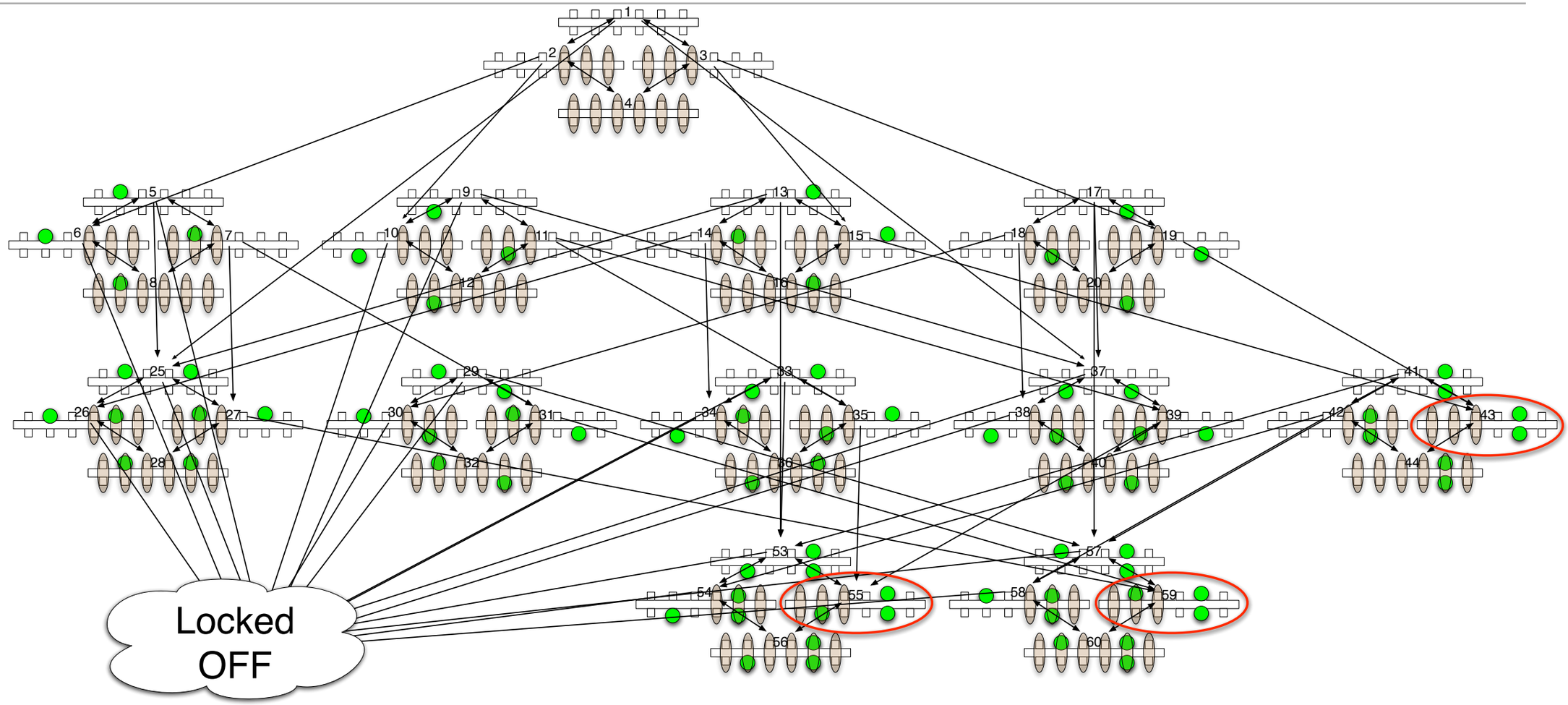
Aggregating Unobservable States



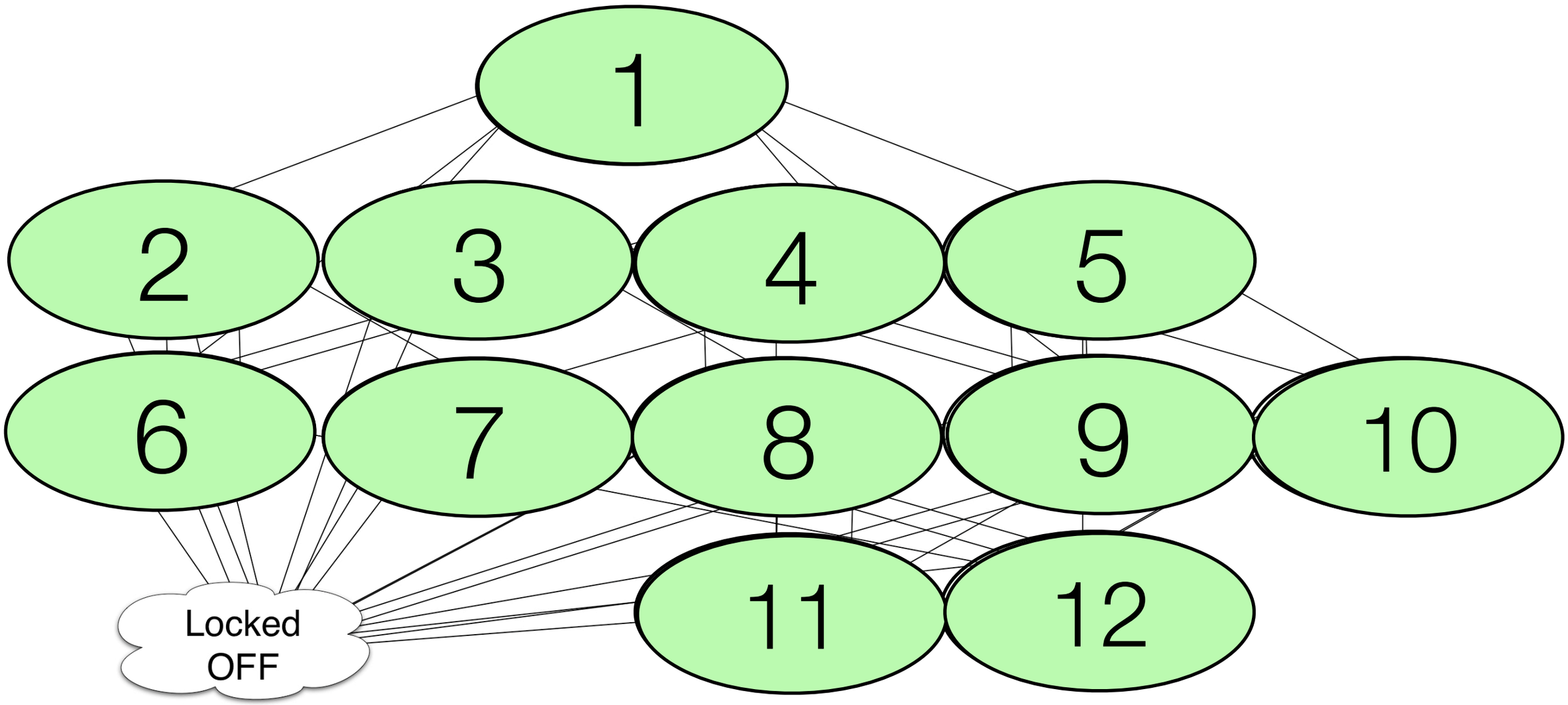
These States are unobservable from the ON states and can be aggregated.

16 Different Methylation Patterns
x 4 Different LRP binding Patterns
=64 Different Operon Configurations!

Aggregating Unobservable States



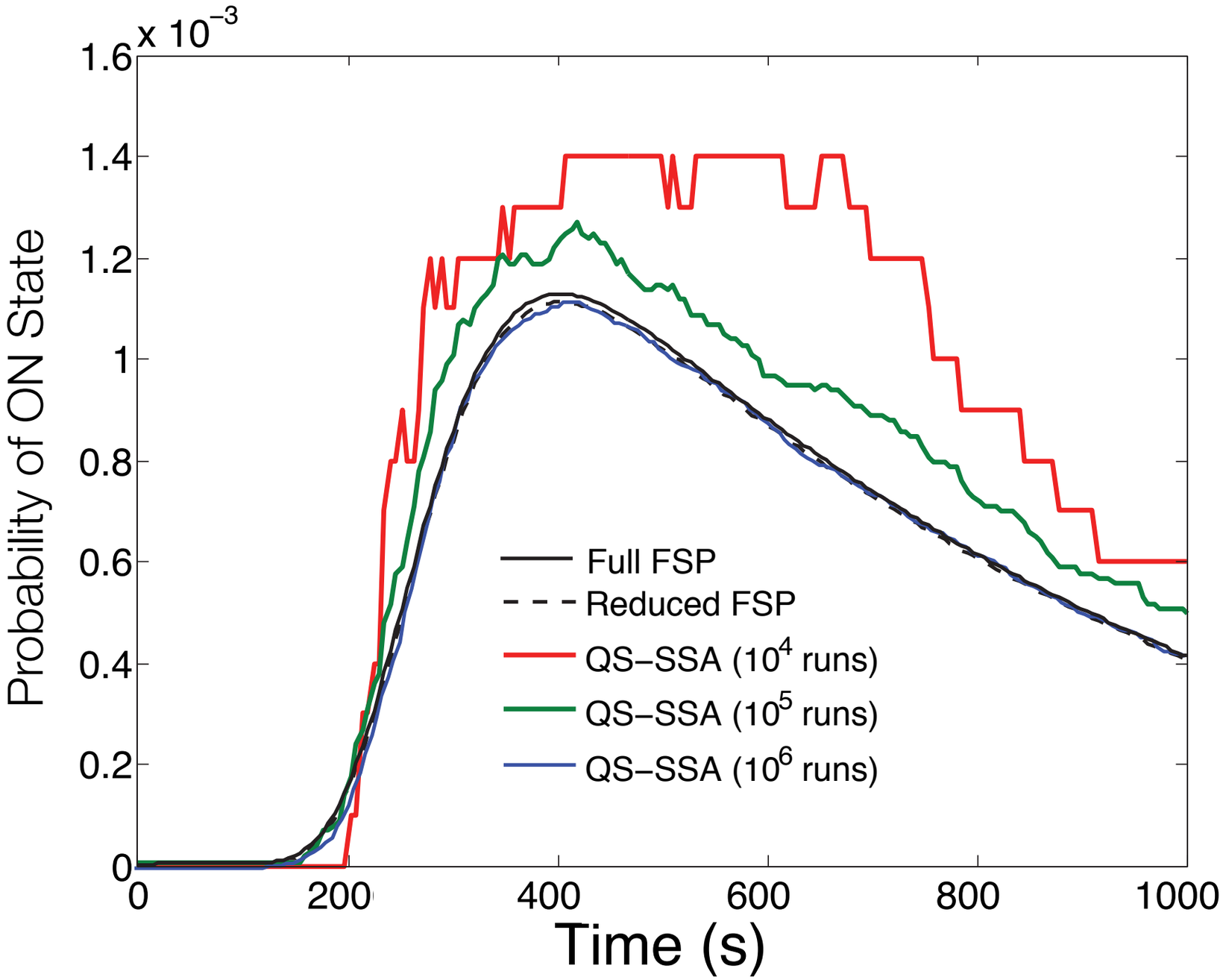
Aggregating Fast States



FSP vs. Monte Carlo Algorithms

Method	# Simulations	Time (s)[a]	Relative Error[b]
Full Model			
FSP	N.A. [c]	42.1	$< 0.013\%$
SSA	10^4	> 150 days	Not available

Comparisons



Conclusions

- Low copy numbers of important cellular components give rise to stochasticity in gene expression. This in turn results in cell-cell variations.

- Organisms use stochasticity to their advantage

- Stochastic modeling and computation is an emerging area in Systems Biology

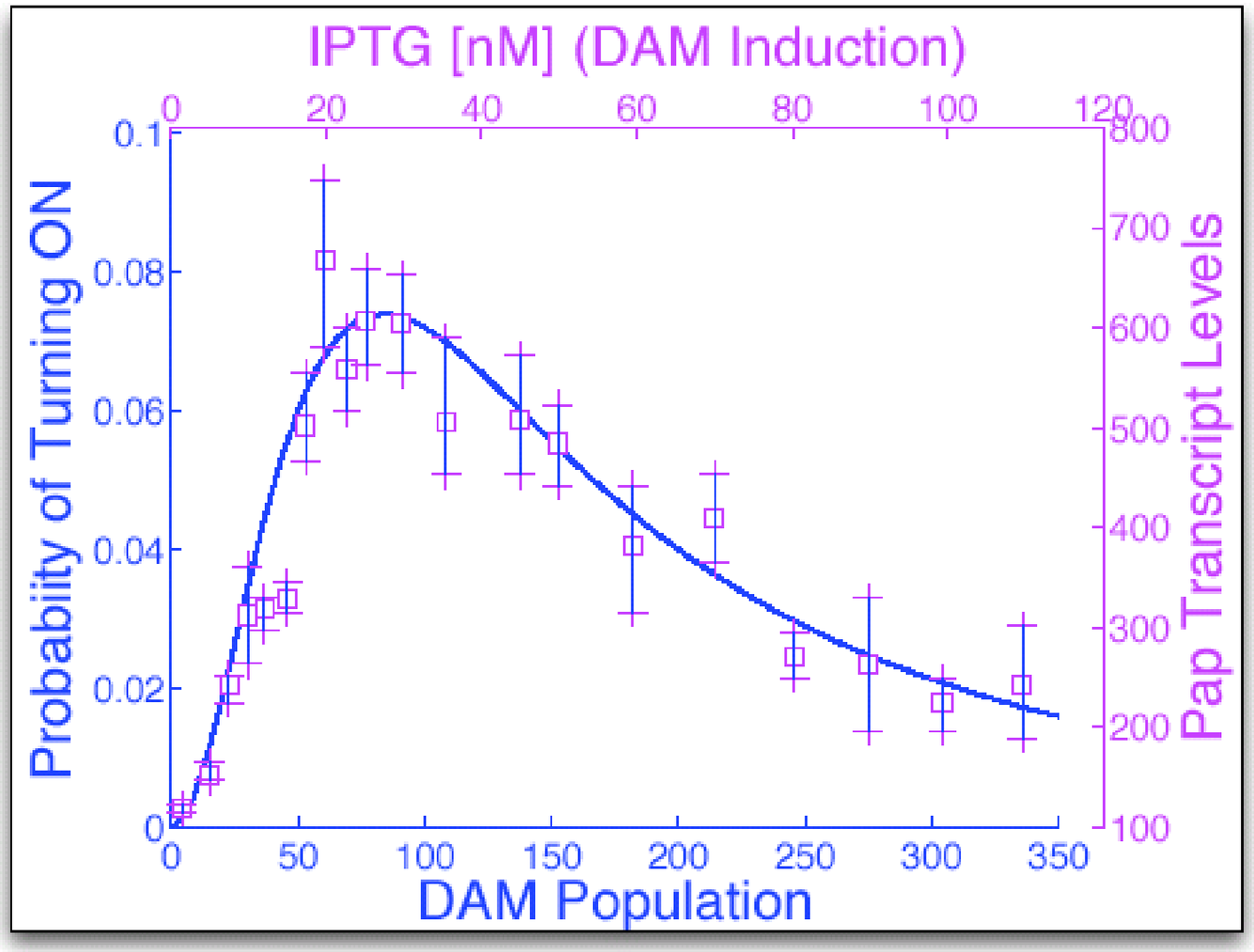
- New tools are being developed. More are needed.

- Many challenges and opportunities for control and system theorists.

Acknowledgments

- Brian Munsky, UCSB
- David Low, UCSB
- Slaven Peles, UCSB

Prediction vs. Experiments



Blue Line -- Predicted probability of expressing Pap versus DAM population.

Magenta -- Experimentally measured Pap transcript levels under different DAM induction levels.

Moment Computations

For the first moment of any X_i, we multiply CME by X_i and take summation over all variables $X_1, ...X_N$.

For the second order moment, $E[X_iX_j]$, we multiply CME by X_iX_j and again take summations:

$$\begin{aligned}
\frac{dE[X_i]}{dt} &= \sum_{k=1}^{M} s_{ik}E[w_k(X)] \\
\frac{dE[X_iX_j]}{dt} &= \sum_{k=1}^{M} (s_{ik}E[X_jw_k(X)] + E[X_iw_k(X)]s_{jk} + s_{ik}s_{jk}E[w_k(X)])
\end{aligned}$$

Let

$$W(X) := [w_1(X) \cdots w_M(X)]^T$$

In matrix notation:

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

The Linear Propensity Case

Suppose that the propensity functions are linear in the states, e.g. $w_k(X) = \alpha_k X_j$ (for some j).

In this case $E[W(X)X^T] = W_X E[XX^T]$ and $E[W(X)] = W_X E[X]$. Then

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

The stationary covariance matrix C is given by the Lyapunov equation

$$S W_X C + C S^T W_X^T + S \ diag(W_X \bar{X}) S^T = 0$$