A humble request to all, Please ask questions or doubts at the end of the presentation. You can note down the slide number given at the bottom-right corner of every slide.

Simulation of Prokaryotic Genetic Circuits

Abhinav Mishra

131503 | 7th Semester

Overview

- Organization of Genetic Regulatory Circuits
- Simulations of Cellular Regulation
- Modelling

Why Simulations are needed ?

- To identify design principles for the biochemically based logic.
- To understand the dynamical response of both normal and mutant cells to environmental and interval signals.
- To predict quantitative effects of mutations on regulatory outcomes.
- To verify consistency and completeness of hypotheses reactions systems.

What are the challenges in Simulations ?

Developing simulation techniques applicable to cellular processes where genetic regulation is centrally important.

- Developmental Differentiation
- Facultative Infection Process
- Cell Cycle Control

In ’61 Cold Spring Harbor Conference on cellular Regulatory Mechanism, regulatory nets are characterized as **‘Circuits’.**

Regulatory circuits

Hierarchical organization

- Regulons – control groups of operons
- Global regulons – multiple pathway regulation (e.g. IHF , σ^{32})
- Often neglected in simulations
- However , needed in some circumstances (e.g. 2 σ factors competing)

Regulatory feedback

- Output influences input signals
- Auto regulatory feedback loops
- In E.coli, there are 107 σ^{70} promoters
  - 68% auto regulating
  - 13% auto activating
- Specialized enzymes often under regulatory control

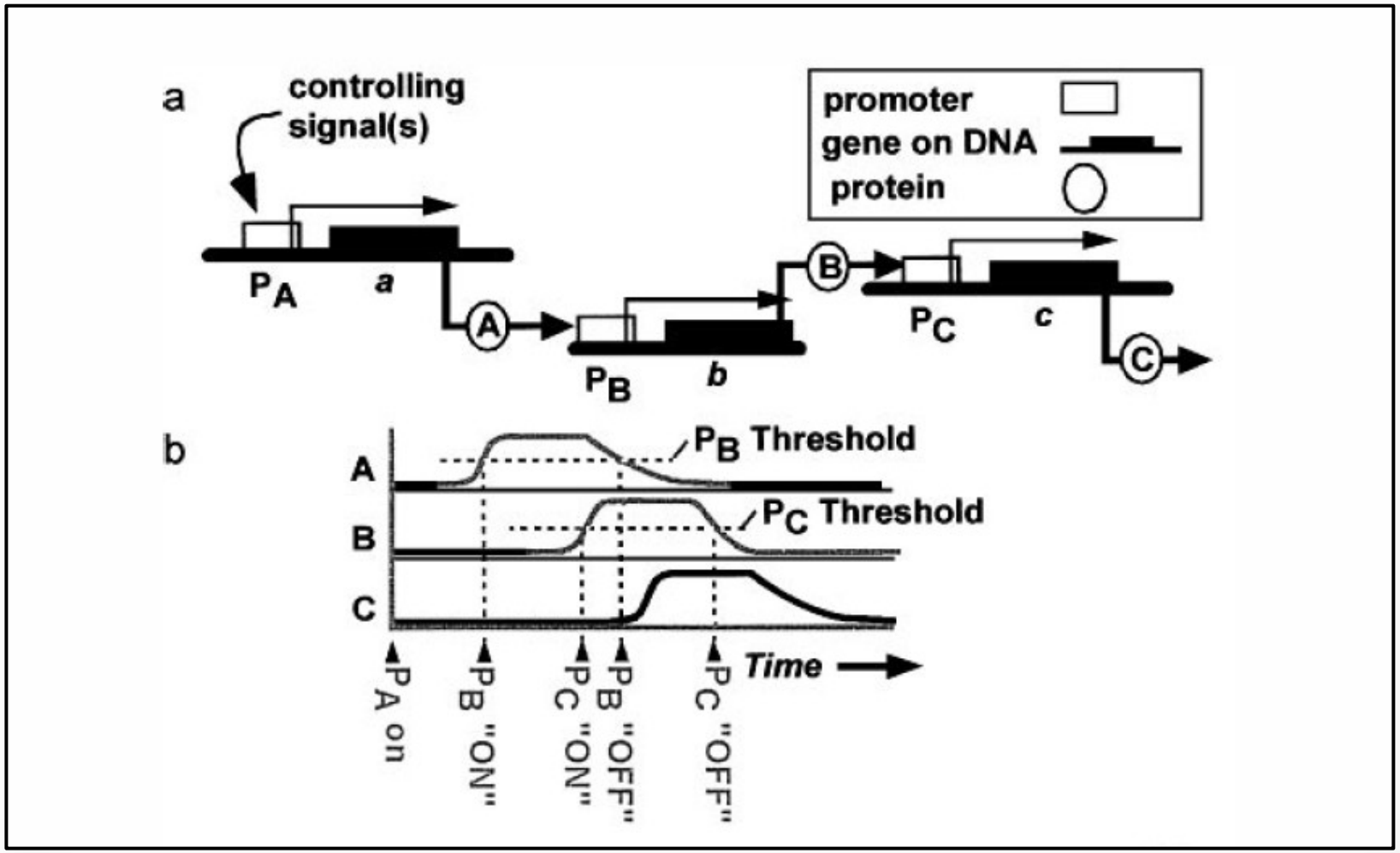
What defines the logic that how well the cell functions at any instant?

Complement of distinct molecules in the cell

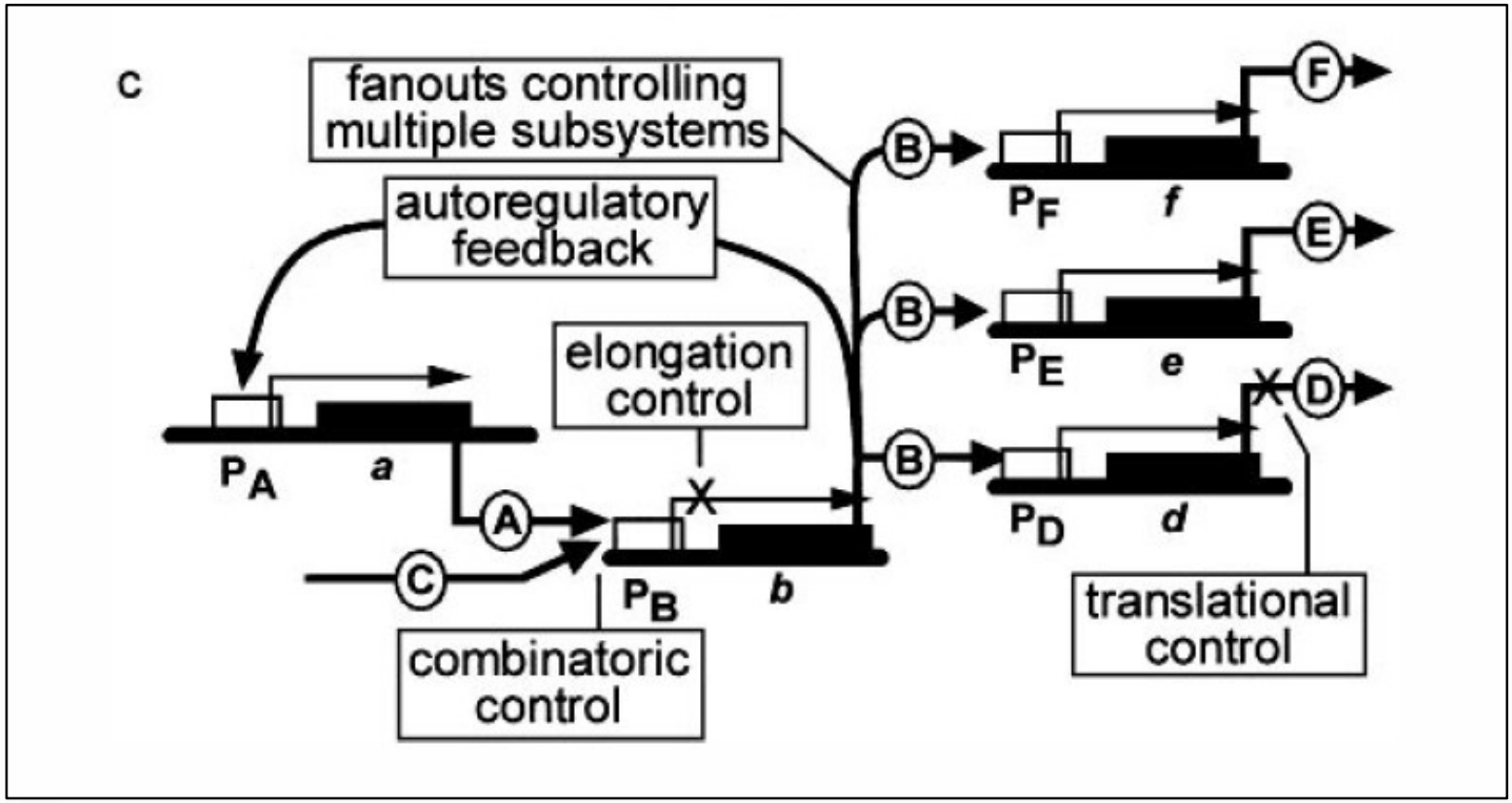
State of DNA | Methylation or Demethylation

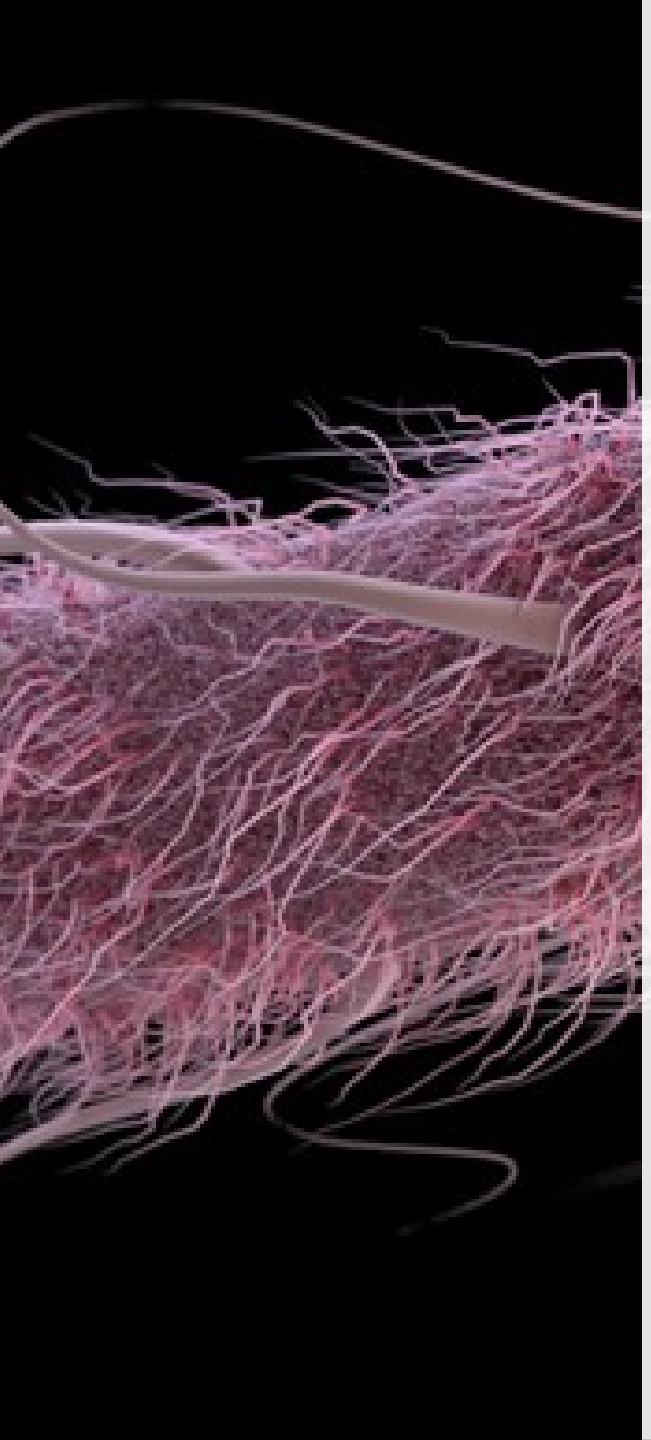
	Electronic logic	Genetic logic
Signals	Electron concentrations	Protein concentrations
Distribution	Point → point (by wires or by electrically encoded addresses)	Point → point (movement by diffusion or active transport by encoded reaction specificity)
Organization	Hierarchical	Hierarchical
Logic type	Digital, clocked sequential logic	Analog unclocked (can approximate asynchronous sequential logic)
Noise	Inherent noise due to discrete electron events and environmental effects	Inherent noise due to discrete chemical reaction events and environmental effects
Signal/noise ratio	Signal/noise ratio high in most circuits	Signal/noise ratio low in most circuits
Switching speed	Fast ($>10^6$ sec^{-1})	Slow ($<10^{-2}$ sec^{-1})

Genetic Cascade



Regulatory mechanisms





Integrating environmental signals

- Chemotactic responses
- Attractant or repellent molecules bind directly to specialized receptors leading to phosphorylation cascade
  - Pulses of agents matched with behavioural changes
  - Mutants shown to have altered enzymatic activity

Cell cycle models

- Genetic regulation coupling to cell cycle
- Modelling of biochemical reactions that support oscillations
  - p^{34}activation, p^{34}/cyclin interactions and cyclin degradation suggested
  - However shown to be far more elaborate

Developmental Switches

- Different physiological states require switching mechanisms
- Cell-density-dependent gene expression
  - Quorum-sensing
    - Higher density = Higher peptide Pheromone concentration
  - Lytic/Lysogenic determination

Modelling

- Promoter control Models
- Stochastic processes in regulatory kinetics
- Modelling macromolecular complexes
- Uncertainty in intracellular environment and reaction rates

Promoter Control Models

$$\dot{x}_i = k_i F_i(x_1, x_2, \ldots, x_n) - k_{di} x_i$$

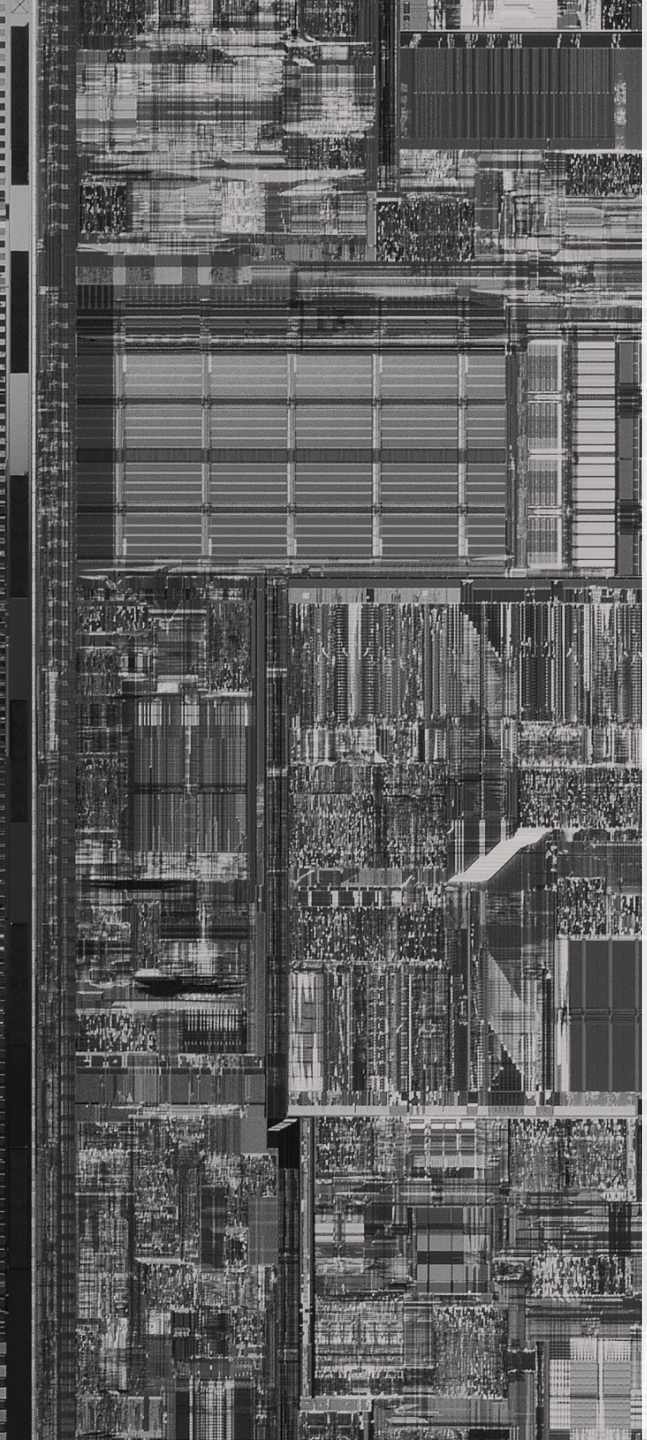
X_i is the concentration of i^{th} protein species

$\dot{X}$ is the time derivative of X_i i.e. $\dot{X} = \frac{dX_i}{dt}$

k_{di} is the degradation rate constant for protein type i

k_i is the rate of protein production with gene type i is ON

F_i is the step function assumed to be 0 or 1 depending on the concentration to threshold values determined by the kinetics of the promoter sites



Assumptions in Boolean Network

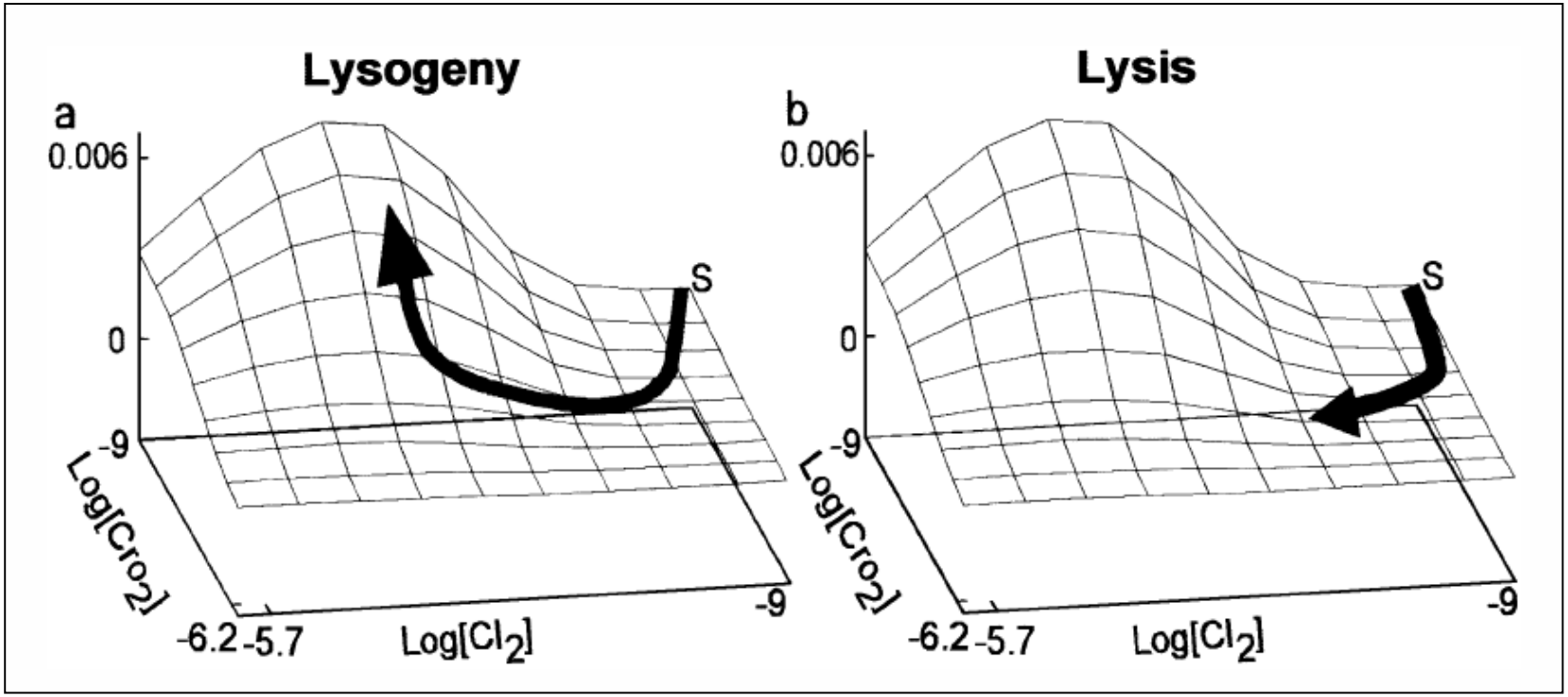
- The state of each gene or other network element can be characterized as either on (one) or off (zero).
- The combinational control of gene expression can be reduced to a “wiring diagram” of the network.
- The computation of the interactions indicated by the wiring diagram can be approximated by Boolean combinational logic rules.
- All elements (to first approximation) update their on or off states synchronously.

Software used: **D.D.L. (Discrete Dynamics Lab)** that computes the behavior of hypothetical networks.

Limitation of Boolean Network

Poor approximation

Shea–Ackers scheme



Bacteriophage λ encodes two repressor proteins

- **Cro repressor** acts to turn off early gene transcription during lytic cycle
- **CI repressor** maintains lysogenic growth

Together, they are known as

Other control mechanism

- Termination sites activation control
- Many post-transcriptional regulations
- Many protein-mediated controls
  - Proteolysis
  - Phosphorylation
  - Methylation

Stochastic Process

- Model macroscopic kinetics of chemical reactions using ordinary differential equations
- Difficult to achieve in genetic reaction due to spatial isolation, low concentration and slow reaction rates
- **Gillespie Algorithm** – calculating the probabilistic outcome of each discrete chemical event
- State vector characterize the states of the system.

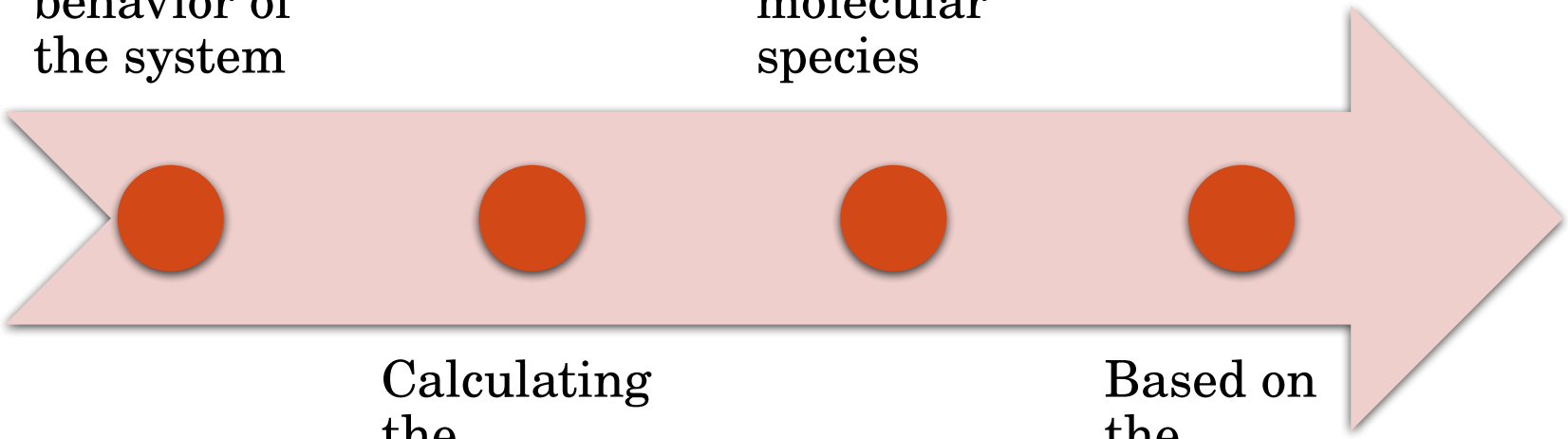
There are two fundamental ways to view coupled systems of chemical equations:

- **Continuous**, represented by differential equations whose variables are concentrations
- **Discrete**, represented by stochastic processes whose variables are numbers of molecules.

Gillespie Algorithm

Stochastic realization of temporal behavior of the system

Resulting changes in the number of each molecular species



Calculating the probabilistic outcome of each discrete chemical event

Based on the application of CME (chemical master equation)

Gillespie Algorithm

Consider, for example,
the set of reactions.
The propensities of the reactions are given by k_1, k_2, …k_5. The constants k_i may be a function of temperature, volume, electrolyte concentration etc.

$$A + B \xrightarrow{k_1} C$$

$$B + C \xrightarrow{k_2} D$$

$$D + E \xrightarrow{k_3} E + F$$

$$F \xrightarrow{k_4} D + G$$

$$E + G \xrightarrow{k_5} A$$

Gillespie Algorithm

Gillespie proposed two exact stochastic simulation algorithms.

Consider a system of r reactions and assume every rate constant k_i are true constants.

At each time step, the system is in exactly one state. A transition consists of executing a reaction so there are at most r possible transitions from a given state.

The key is to **choose random numbers** using a computer random number generator and to use those random numbers to **pick transitions**.

Gillespie Algorithm

Direct Method

Calculates explicitly which reaction occurs next and when it occurs

First Reaction Method

Generates for each reaction μ a putative time τ_μ at which reaction μ occurs, then chooses the reaction μ^* with the smallest time $\tau_\mu{}^*$ (the first reaction) and executes reaction μ^* at time $\tau_\mu{}^*$

Gillespie Algorithm : Direct Method

Probability Density Function

$$P(\mu,\tau)d\tau = a_\mu \exp(-\tau \sum_j a_j)\, d\tau \qquad (1)$$

$$\int_0^\infty \boldsymbol{P(\mu,\tau)d\tau = a_\mu \sum_j a_j} \qquad (2)$$

$$\sum_J {}^{\mu}\boldsymbol{P(\mu,\tau)d\tau = \sum_j a \exp(-\tau \sum_j a)\, d\tau} \qquad (3)$$

Gillespie Algorithm : Direct Method

1. **Initialize** (initialize numbers of molecules, set $t \leftarrow 0$).
2. **Calculate** the propensity function, a_i, for all i.
3. **Choose** μ according to the distribution in eq 2.
4. **Choose** τ according to an exponential with parameter $\sum_j aj$ (as in eq 3).
5. **Change** the number of molecules to reflect execution of reaction μ. Set $t \leftarrow t + \tau$.
6. Go to Step 2.

Gillespie Algorithm : First Reaction Method

Generates a **putative time τ_i** for each reaction to occur - a time the reaction would occur if no other reaction occurred first - then lets μ be the reaction whose putative time is first, and lets τ be the putative time τ_μ.

The algorithm of the previous subsection is direct in the sense that it ***generates μ and τ directly.***

Gillespie Algorithm : First Reaction Method

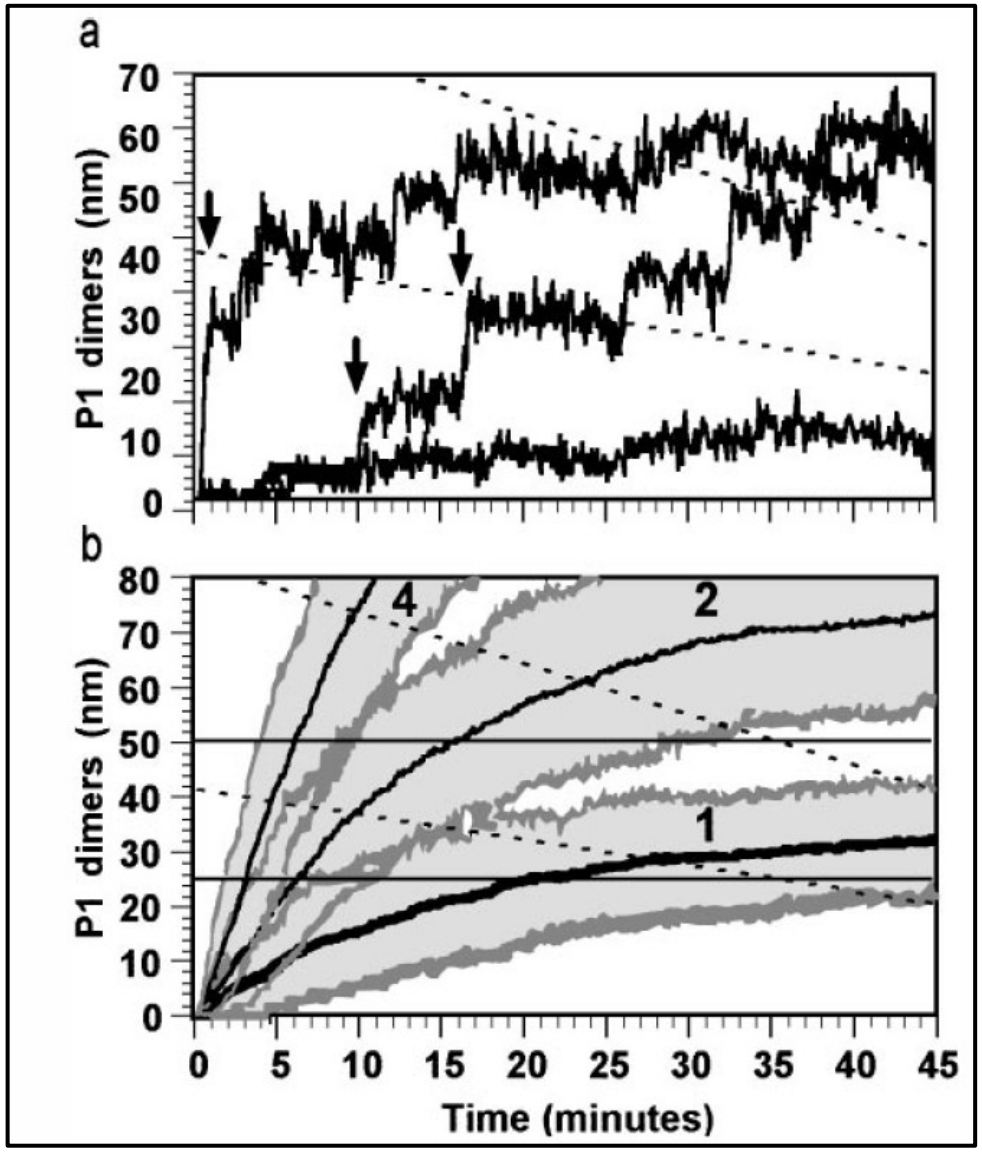
1. **Initialize** (i.e., set initial numbers of molecules, set $t \leftarrow 0$).
2. **Calculate** the propensity function, a_i, for all i.
3. For each i, **generate** a putative time, τ_i, according to an exponential distribution with parameter a_i.
4. Let μ be the reaction whose putative time, τ_μ, is least.
5. Let τ be τ_μ.
6. **Change** the number of molecules to reflect execution of reaction μ. Set $t \leftarrow t + \tau$.
7. Go to Step 2.

Gillespie Algorithm : First Reaction Method

This algorithm **uses *r* random numbers** per iteration (where r is the number of reactions), takes time proportional to r to update the a_i s, and takes time proportional to r to identify the smallest τ_μ.

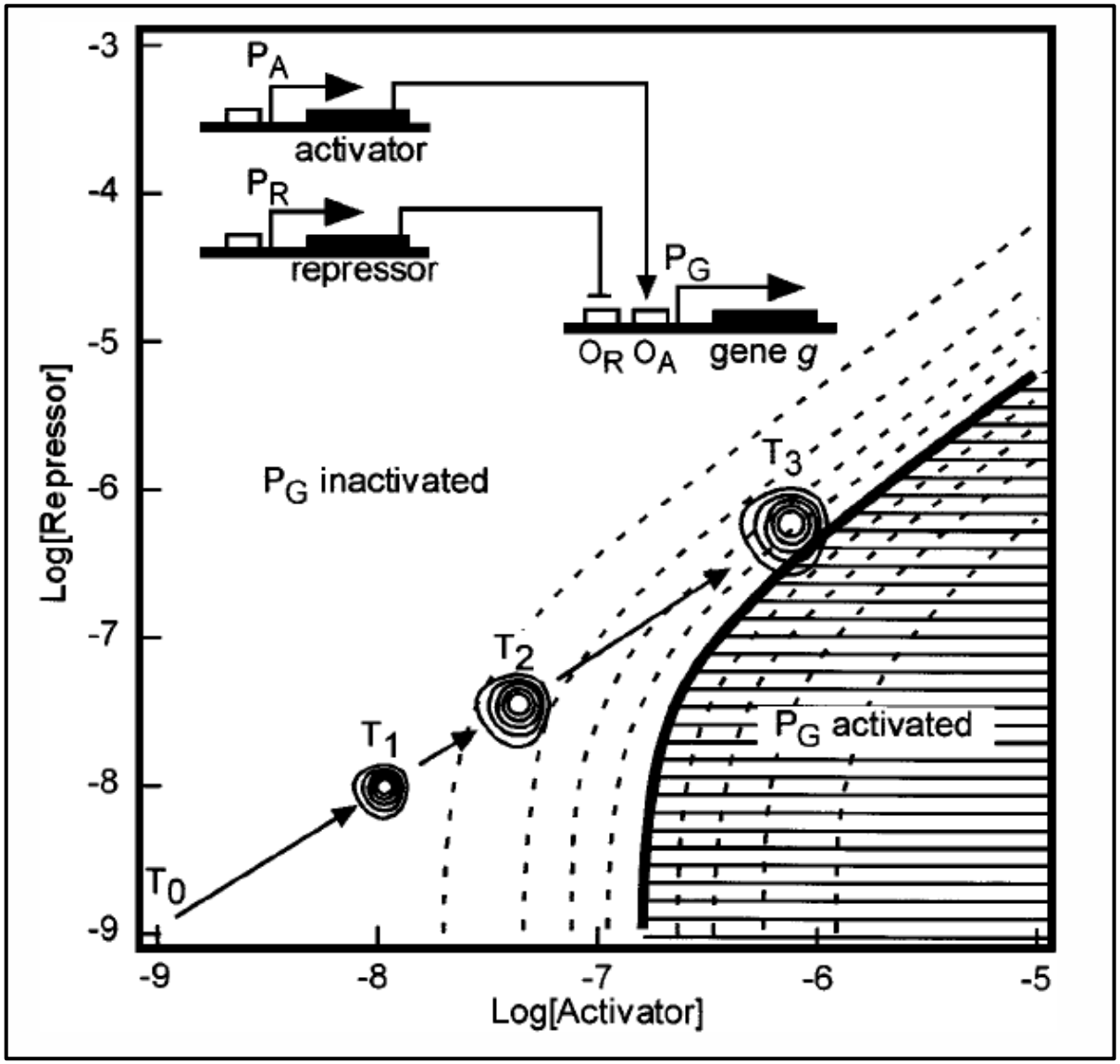
Stochastic Process

- Random burst of numbers of protein
- Timing uncertainty
- Stronger promoter
- Higher gene dosage
- Lower signal threshold



Stochastic Process

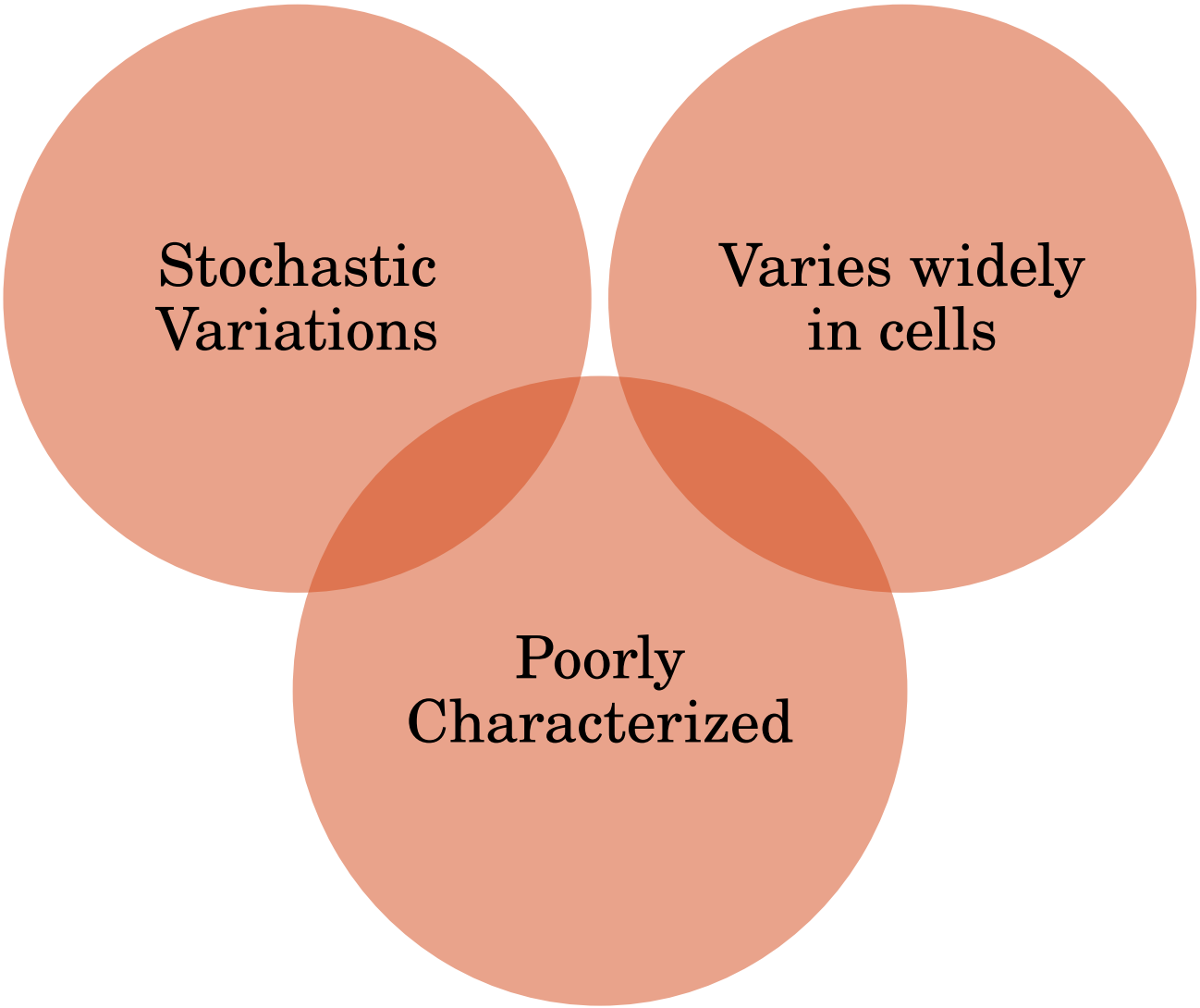
Different Activation time due to variation of the concentration



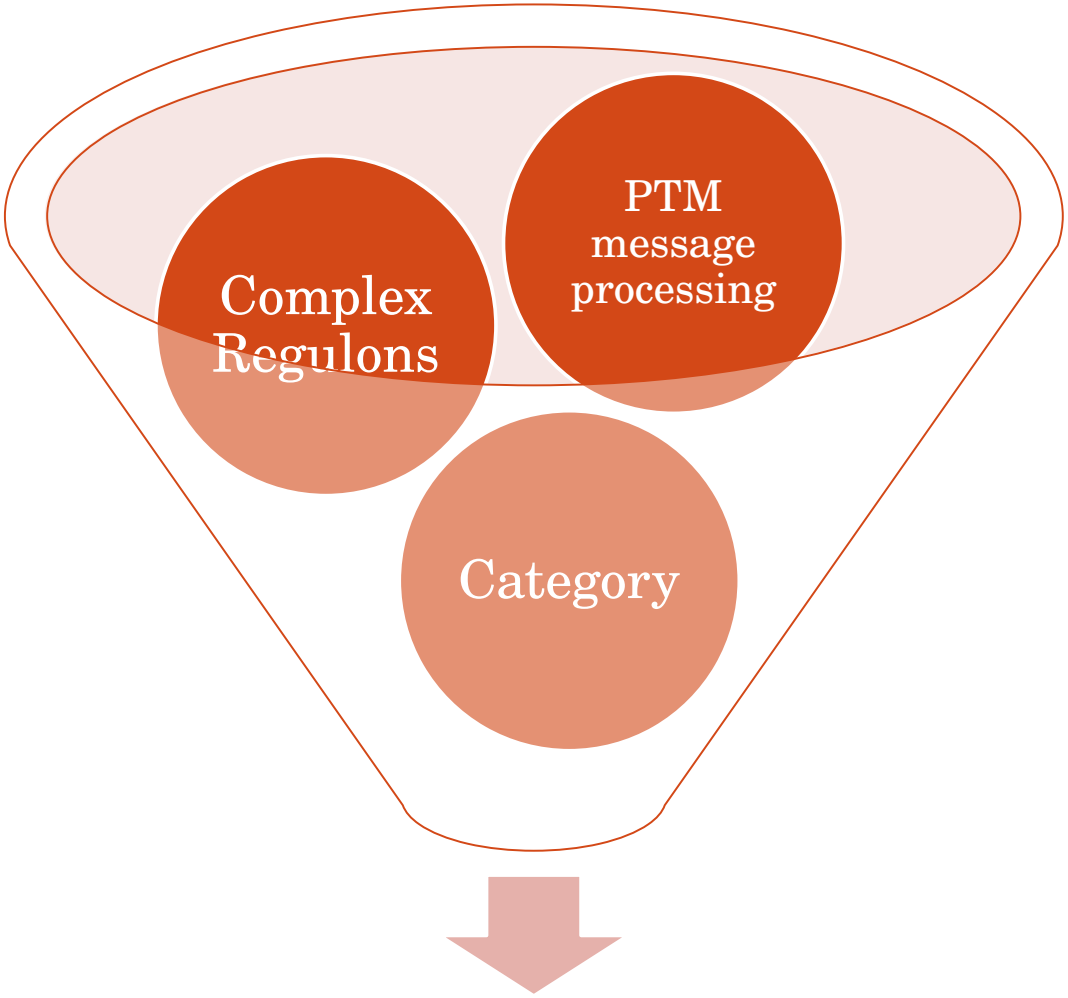
Modelling macromolecular complex

- Realistic Modelling | Central Challenge
- Complicated network mechanism
- Dynamic behaviour

Uncertainty of Intracellular Reaction Features



Challenges



Eukaryotic Networks

Opportunities

1

- To recognize Common Circuit Motifs.

2

- To identify function of individual protein in regulatory control mechanisms.

3

- To redesign circuit for altered functions.

References

- **McAdams, H. H. and A. Arkin (1998). "Simulation of prokaryotic genetic circuits." Annu Rev Biophys Biomol Struct 27: 199-224.**
- Gillespie, D. T. (1992). "A rigorous derivation of the chemical master equation." Physica A: Statistical Mechanics and its Applications **188**(1): 404-425.
- Gibson, M. A. and J. Bruck (2000). "Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels." The Journal of Physical Chemistry A **104**(9): 1876-1889.

Further Readings

- Smadbeck, P. and Y. N. Kaznessis (2014). "Solution of Chemical Master Equations for Nonlinear Stochastic Reaction Networks." Curr Opin Chem Eng **5**: 90-95.
- Gillespie, D. T. (1976). "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions." Journal of Computational Physics **22**(4): 403-434.
- Gillespie, D. T. (1977). "Exact stochastic simulation of coupled chemical reactions." The Journal of Physical Chemistry **81**(25): 2340-2361.
- Miller, M. B. and B. L. Bassler (2001). "Quorum sensing in bacteria." Annu Rev Microbiol **55**: 165-199.
- Shea, M. A. and G. K. Ackers (1985). "The OR control system of bacteriophage lambda. A physical-chemical model for gene regulation." J Mol Biol **181**(2): 211-230.

- Brophy, J. A. N. and C. A. Voigt (2014). "Principles of genetic circuit design." Nat Meth **11**(5): 508-520.
- Gillespie, D. T. (1976), 'A general method for numerically simulating the stochastic time evolution of coupled chemical reactions', *Journal of Computational Physics* **22** (4), 403-434.
- Gillespie, D. T. (2007). "Stochastic simulation of chemical kinetics." Annu Rev Phys Chem **58**: 35-55.
- Li, H., et al. (2008). "Algorithms and software for stochastic simulation of biochemical reacting systems." Biotechnol Prog **24**(1): 56-61.
- Hyver, C. and H. Le Guyader (1990). "MPF and cyclin: modelling of the cell cycle minimum oscillator." Biosystems **24**(2): 85-90.

$$\frac{\partial}{\partial t}$$

Thank You.

$$\nabla$$