Stochastic Processes in Gene Regulatory Circuits (II)

Kyung Hyuk Kim

Systems and Synthetic Biology, 498A

Overview

- Modeling of stochastic processes.
  - Assumptions
  - Reaction rate equations: Deterministic vs. stochastic
- When to implement stochastic models
- Gillespie Algorithm

The Origin of Stochasticity

Stochasticity:

- Thermal motion: Intracellular macro-molecules collide with each other and with water molecules.
- At a time scale of 10^{-6} sec, velocity information of typical proteins is lost due to the random collisions.
- Interactions between water molecules and macromolecules are often modeled with white noise and friction force.

Homogeneity:

- At a time scale of 1 msec, typical small proteins floating in the cytoplasm of E. coli will diffuse throughout the cell. Position information of the proteins is therefore quickly lost due to rapid diffusion.
- Non-specific binding of TF's to DNA strands ~ 5msec (lifetime).
- At a time scale of minutes, we assume that TF's are uniformly distributed in cytoplasm.

Model of the Homogeneous System

- The state of a Homogeneous System is described by numbers of molecules of each species.
- Reactions occur randomly with a probability given by rate constants and the number of combinatory cases.

- No force fields
- No velocities
- No spatial variation
- Markov process : all the previous information is lost and current information determines the next future events.

- Very idealized systems!

N_{red} , N_{blue}

Rate Constants

Stochastic Reaction Systems

When the numbers of molecules (n) are small:

- Concentrations fluctuate in a discrete manner.

- Concentrations fluctuate significantly.

Discrete Noise for Small n

- The number of *LacI* tetrameric repressor protein in E.coli ~ 10 molecules.
- If one *LacI* repressor binds to a promoter region, the number of free *LacI* repressors = 9.
- 10% change in its concentration and number! The change becomes discrete.

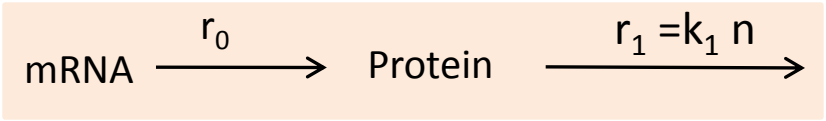
Significant Noise for Small n

- Consider proteins that are translated from mRNA and degrade.

r_0 is assumed to be constant.

r_1 is proportional to the number of protein molecules, n.

Poisson Distributions in Steady States

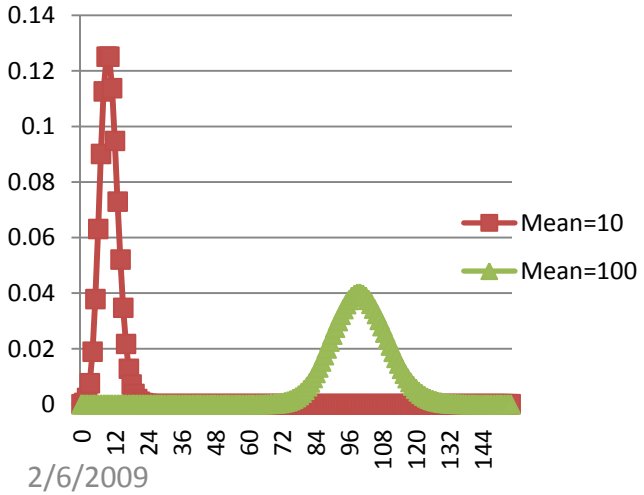


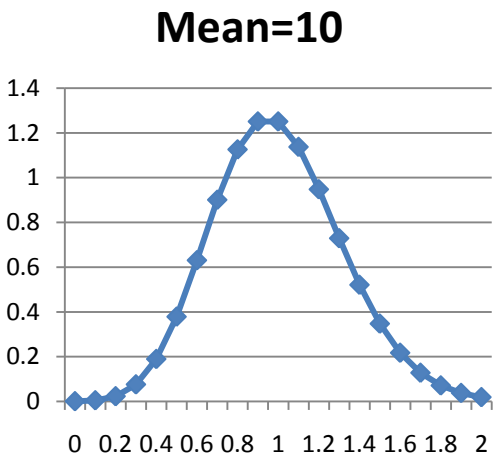
The stationary statistics of n follows a Poisson distribution.

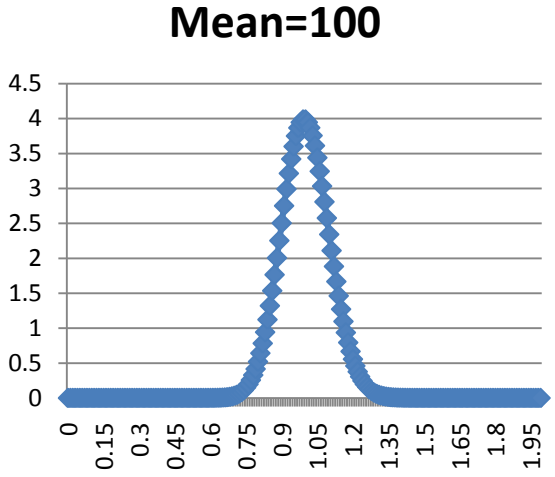
The variance of n [Var(n)] = The mean of n [Mean(n)].

Standard deviation of n [Std(n)] = $[\text{Var}(n)]^{1/2} = [\text{Mean}(n)]^{1/2}$.

As the Mean(n) gets larger, the distribution becomes narrower in the relative terms.







Reaction Rate Equations

- $X + Y \rightarrow Z \qquad v = k_1[X][Y].$
- $Z \rightarrow W \qquad v = k_1[Z].$
- $X + X \rightarrow X_2 \qquad v = k_2[X]([X] - 1],$

When the number of reactant X is 4, $v=k_2$ 4*3.

Gillespie Algorithm

Exact simulation algorithm for chemical master equations.

At each time point we must answer the following two questions:

1. Determine **when** the next reaction will occur.

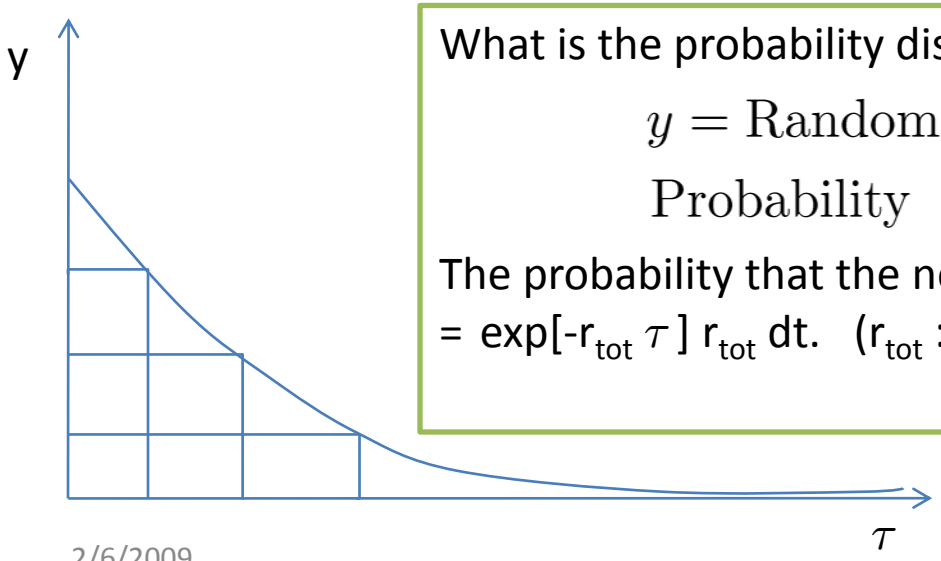
2. Determine **which** reaction will occur.

Gillespie Algorithm

1. Determine **when** the next reaction will occur.

Compute the time of next reaction:

$$\tau = \frac{-\ln(\text{Random Number in } (0,1])}{r_{tot}} \qquad r_{tot} = r_0 + r_1$$



What is the probability distribution of τ?

$$y = \text{Random Number in } (0,1] = e^{-r_{tot}\tau}$$

$$\text{Probability} \ \propto \ |\text{Slope of } e^{-r_{tot}\tau}| = r_{tot}e^{-r_{tot}\tau}$$

The probability that the next reaction will occur at [τ , τ + dt]
= exp[-r_{tot} τ] r_{tot} dt. (r_{tot} : normalization factor)

Gillespie Algorithm

2. Determine **which** reaction will occur.

The probability to choose a reaction r_i among all possible reactions

$$= \frac{r_i}{r_{tot}}$$

Therefore,

The probability that next reaction r_0 will occur during

$$[\tau, \tau + dt] = e^{-r_{tot}\,\tau}\; r_{tot}\; dt\; \frac{r_o}{r_{tot}}$$

$$= e^{-r_{tot}\,\tau}\; r_o\; dt$$

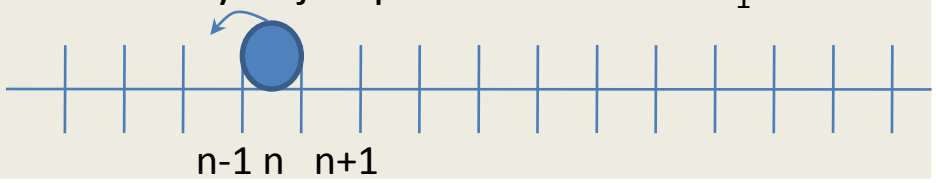
Alternative Simulation Approach

mRNA $\xrightarrow{r_0}$ Protein (n) $\xrightarrow{r_1 = k_1 n}$

The number of proteins are updated at each time step $t_m = m\ dt$.

Probability to jump from n to n+1 = $r_0\ dt$

Probability to jump from n to n-1 = $k_1\ n\ dt$



Probability to stay = $[1- (k_1\ n + r_0)\ dt\ ]$

1. At t=0, start from n=10.
2. Generate a random number "*rand*" between (0, 1].
3. If *rand* < $r_0\ dt$, n=11.
4. Else if *rand* < $(r_0 + k_1\ n)\ dt$, n=9.
5. Else if n=10.
6. This new n is the number of protein at t_1=dt.
7. Repeat step 3-6 and t_2=2dt.
8. …..

This can be less efficient than the Gillespie algorithm. It is possible that reactions do not occur.

Alternative Simulation Approach

- Initially, the number of proteins = n.
- What is the probability that no reaction occurs up to time τ = m dt?

$$\left(1-(k_1 n + r_0)dt\right)^m = e^{m\log\left(1-(k_1 n+r_0)dt\right)} \simeq e^{-m(k_1 n+r_0)dt} = e^{-(k_1 n+r_0)\tau}$$

The probability that next reaction r_0 will occur during $[\tau, \tau + dt]$ is

$$e^{-(k_1 n+r_o)\,\tau}\, r_o\, dt$$

Two different simulation approaches

- Gillespie
  - Prob[No reaction occurs until τ and any single reaction occurs between τ and τ + dt] *

    Prob[a specific reaction (i-th) should have occurred]

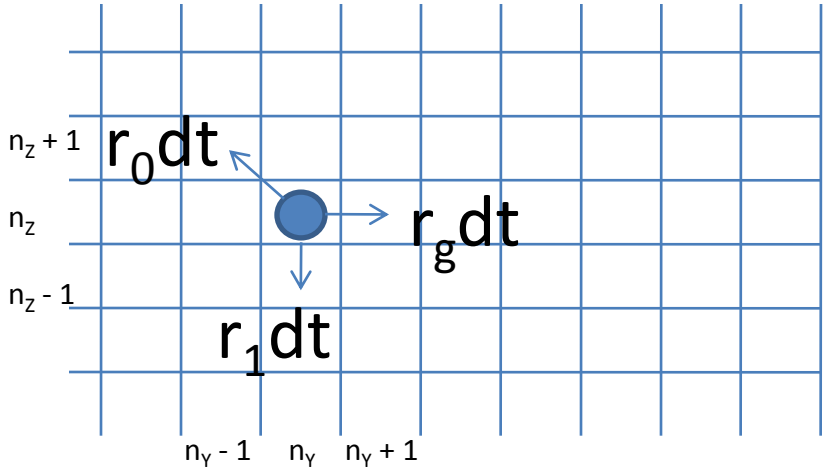
- Alternative approach
  - Prob[No reaction occurs until τ] *

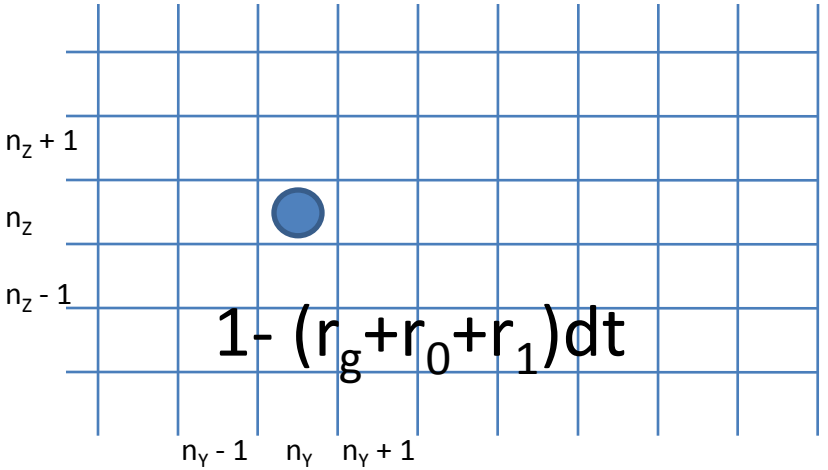
    Prob[a specific reaction (i-th) occurs between τ and τ + dt]

Example 2

$$X \xrightarrow{r_g} Y \ (n_Y) \xrightarrow{r_0 = k_0 n_Y} Z \ (n_Z) \xrightarrow{r_1 = k_1 n_z}$$

- State: "n" → "(n_Y , n_Z)"





- The probability that next reaction r_0 will occur during $[\tau , \tau + dt]$ is

$$e^{-(r_g + r_o + r_1)\tau} r_o dt$$

Gillespie Algorithm

Exact simulation algorithm for chemical master equations.

At each time point we must answer the following two questions:

1. Determine **when** the next reaction will occur.

$$\tau = \frac{-\ln(\text{Random Number in } (0,1])}{r_{tot}} \qquad r_{tot} = \sum_i r_i$$

2. Determine **which** reaction will occur.

$$\text{Probability to choose a reaction } r_i = \frac{r_i}{r_{tot}}$$

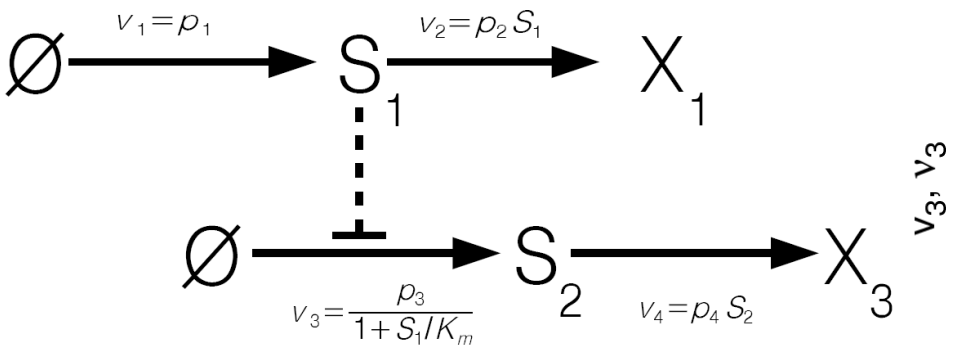
Stochastic Phenomena

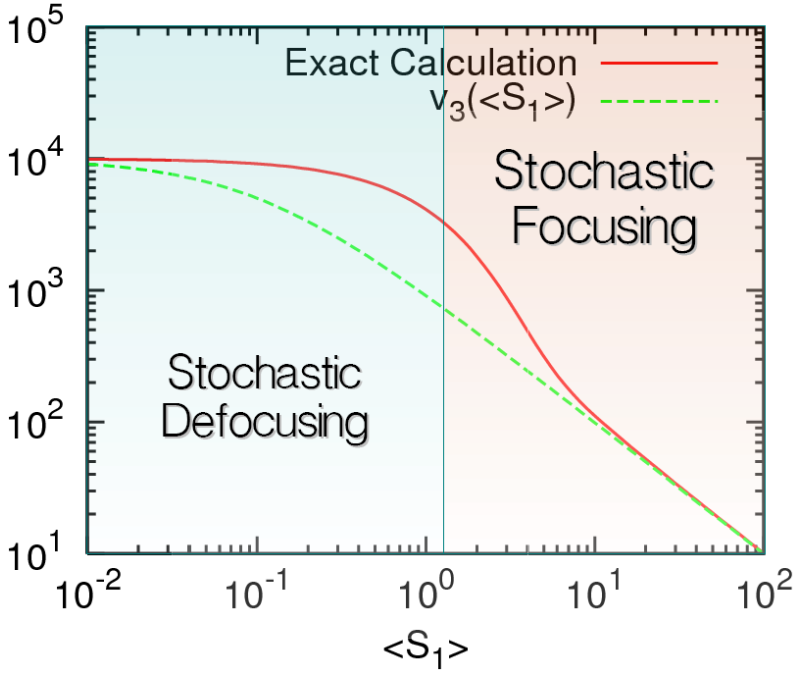
- Stochastic Focusing
- Stochastic Switching
- Single Events (λ Phage)
- Multiplicative Noise Effects

Stochastic Focusing

- Stochastic Focusing:
  Sensitivity increase due to stochastic effects.

  [Paulsson, et al. PNAS 97, 7148-7153 (2000)]
- Two step cascade reactions:





Stochastic Focusing

- Stochastic Focusing:

  Sensitivity increase due to stochastic effects.

- Sensitivity:

$$\text{Sensitivity} = \frac{\%\ \text{Change of Response Signal}}{\%\ \text{Change of Source Signal}}$$

The sensitivity can be used to estimate how a system responds due to changes in the environment.

$$\text{Sensitivity} = \frac{dX}{dY}\frac{Y}{X} = \frac{d\ln X}{d\ln Y}.$$

where we used $\dfrac{d\ln f(x)}{dx} = \dfrac{1}{f(x)}\dfrac{df(x)}{dx}, \quad d\ln f(x) = \dfrac{1}{f(x)}df(x).$

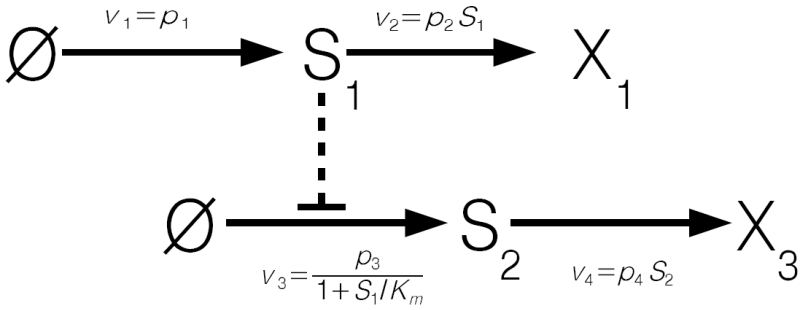
Stochastic Focusing

- Stochastic Focusing:

  Sensitivity increase due to stochastic effects.

  [Paulsson, et al. PNAS 97, 7148-7153 (2000)]

  Two step cascade reactions



Source Signal = S1

Response Signal = S2

$$\text{Sensitivity} = \frac{d\langle S_2 \rangle}{d\langle S_1 \rangle}\frac{\langle S_1 \rangle}{\langle S_2 \rangle} = \frac{d\ln\langle S_2 \rangle}{d\ln\langle S_1 \rangle}$$

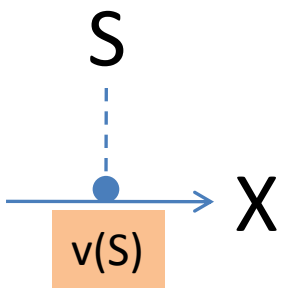
$$\langle x \rangle = Mean(x)$$

Stochastic Focusing

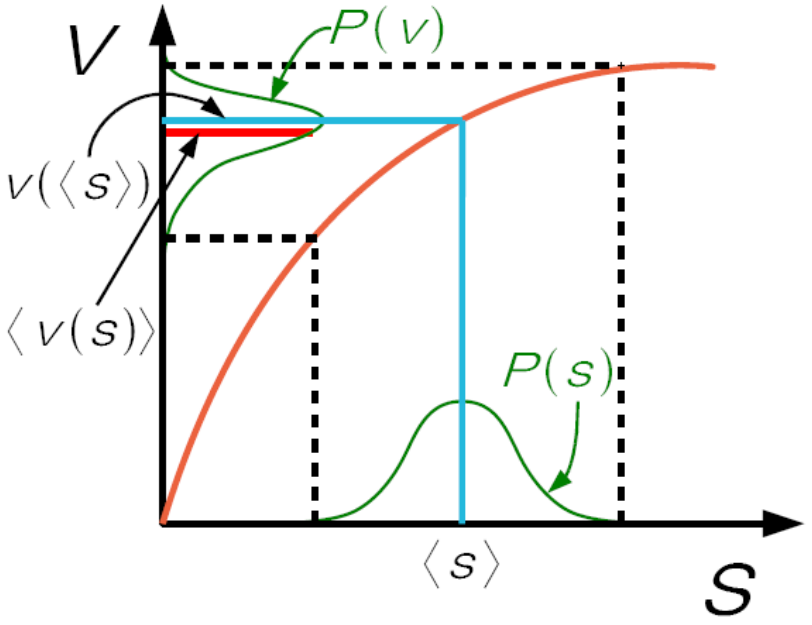
Fluctuations in the concentration of S leads to fluctuations in the reaction rate v(S).

How does the mean rate of reaction change with the noise?

E.g.,



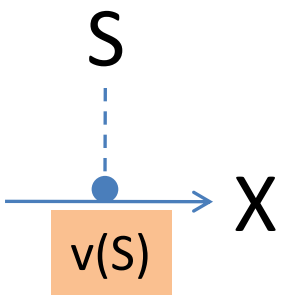
$$v(S) = \frac{k_2 S}{K_M + S}$$

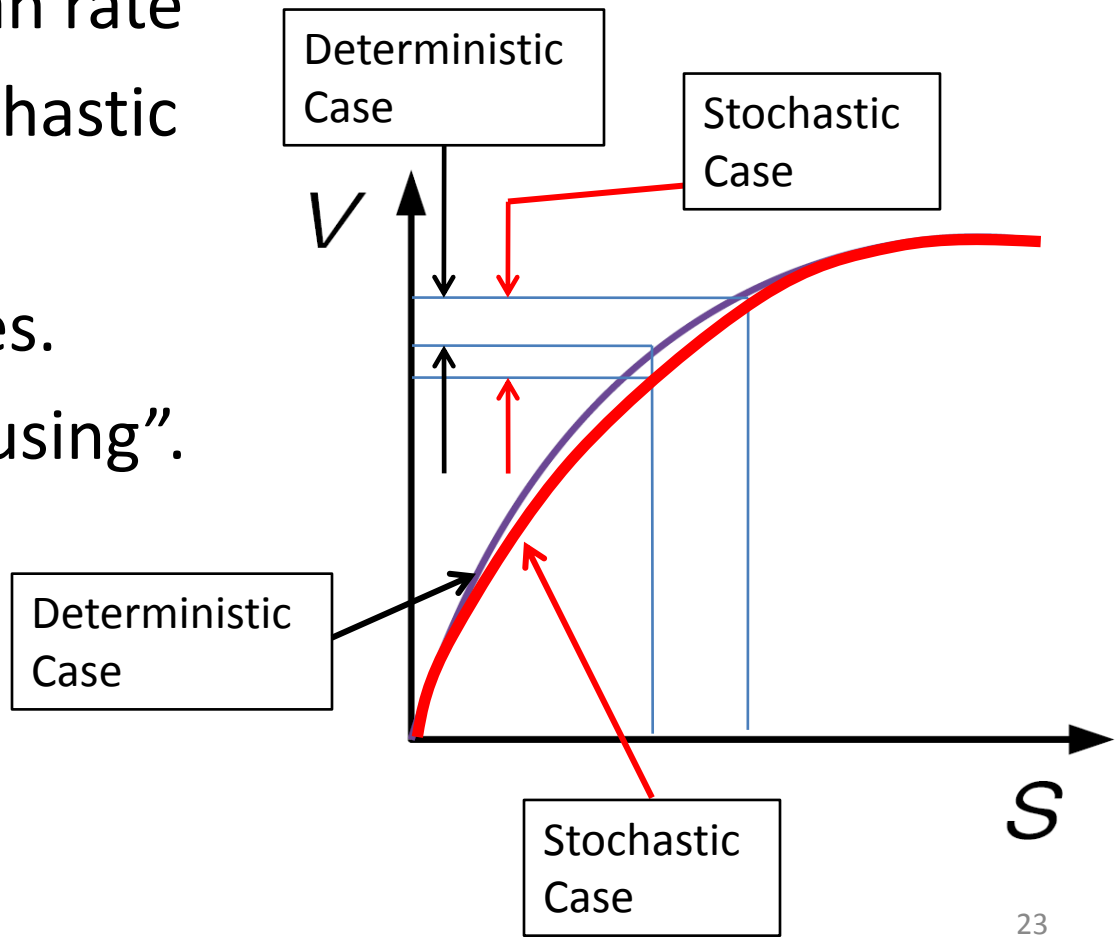


Stochastic Focusing

- We can change mean S.
- The change of mean rate increases with stochastic noise.
- Sensitivity increases.

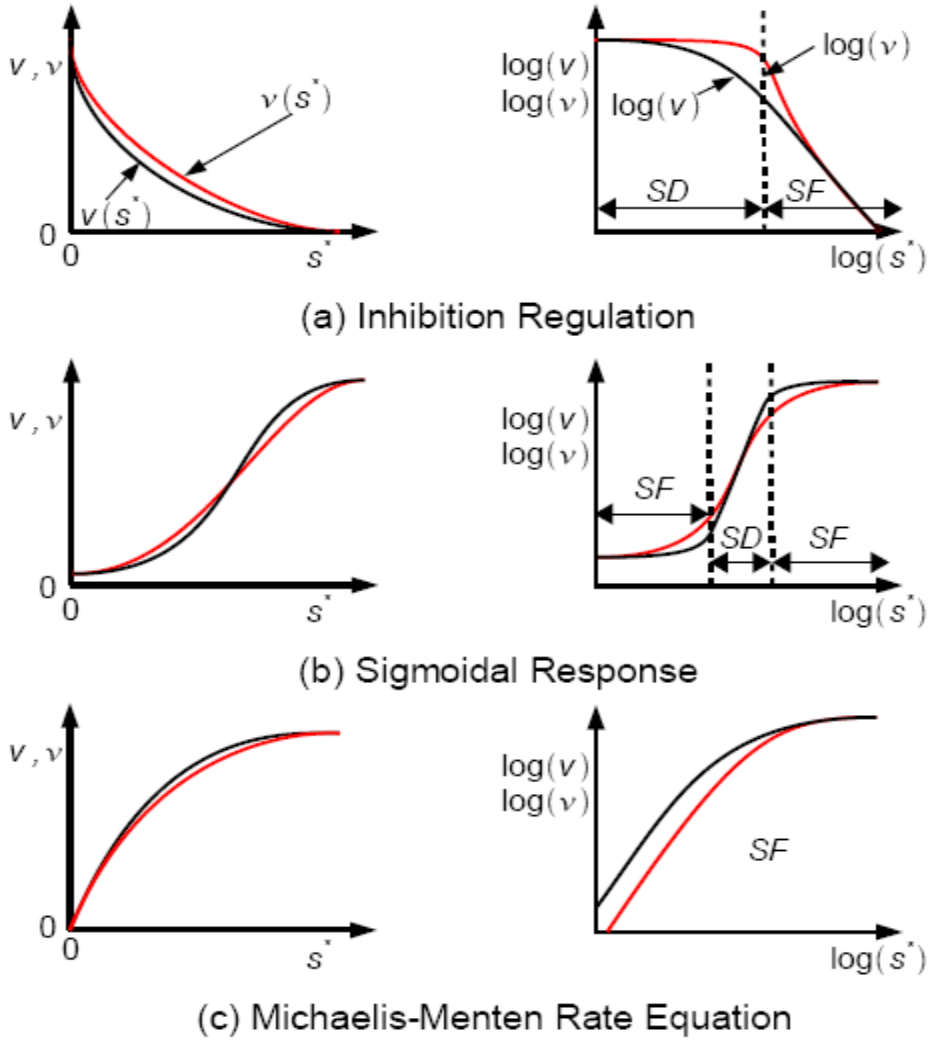
  → “Stochastic Focusing”.





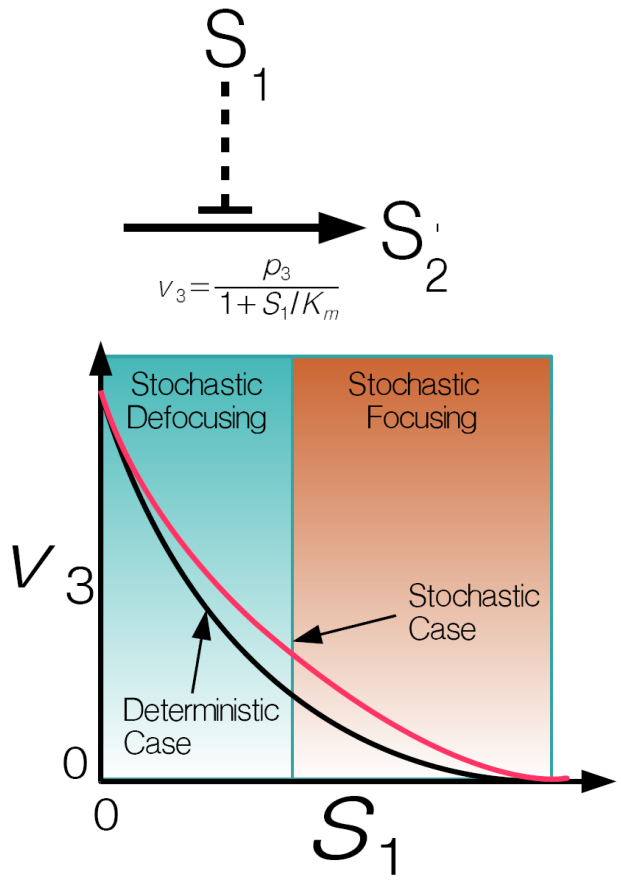
Stochastic Focusing-Defocusing Compensation

- Stochastic focusing can occur in one region of the curve and stochastic defocusing in another region.

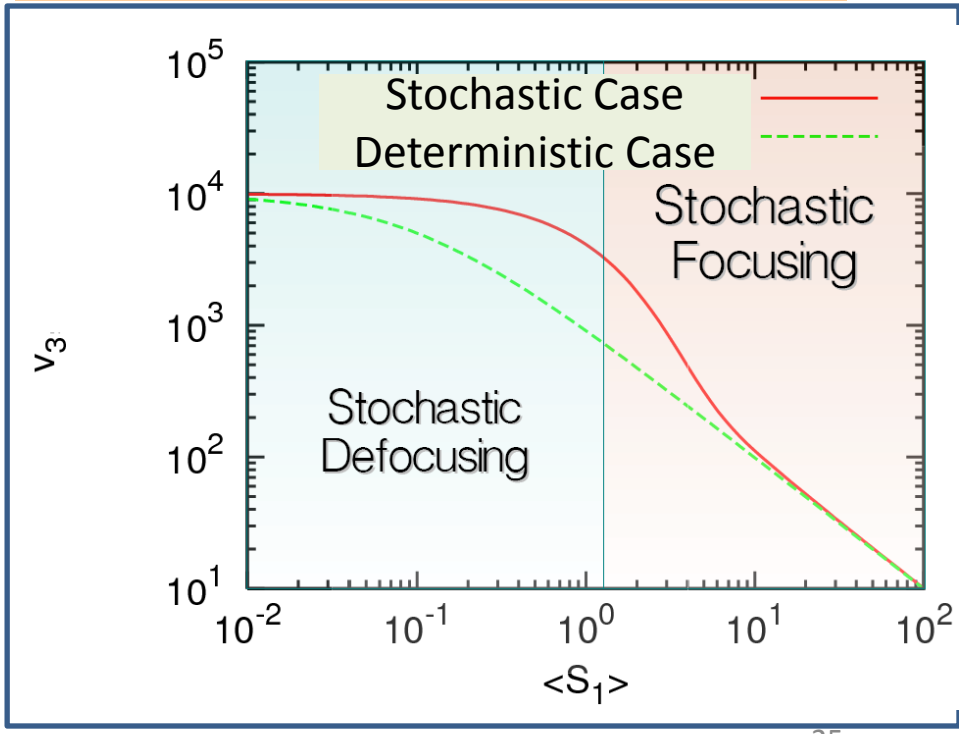


Stochastic Focusing

Two step cascade reactions [Paulsson, et al. PNAS 97, 7148 (2000)]

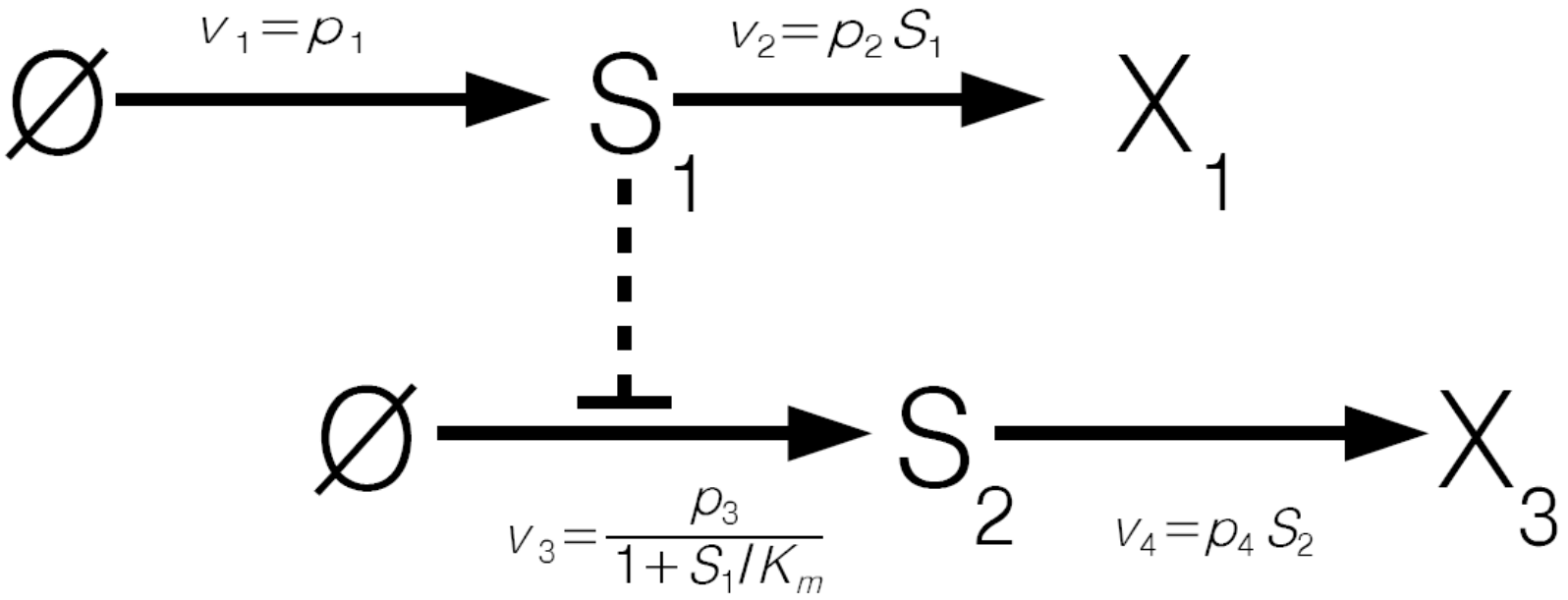


$$\text{Sensitivity} = \frac{d\langle v_3\rangle}{d\langle S_1\rangle}\frac{\langle S_1\rangle}{\langle v_3\rangle} = \frac{d\ln\langle v_3\rangle}{d\ln\langle S_1\rangle}.$$



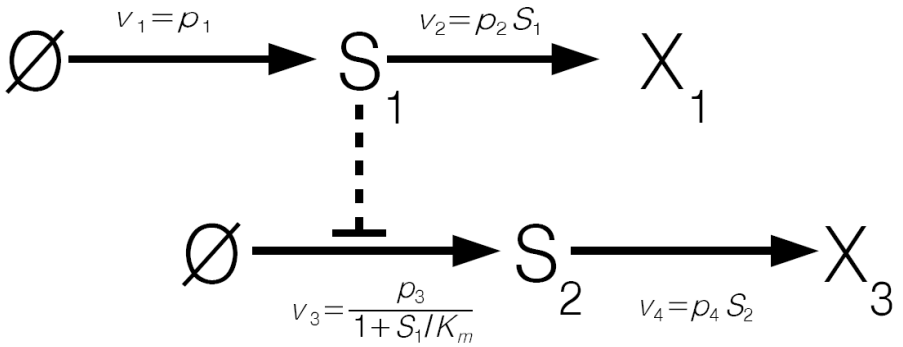
Stochastic Focusing

Consider these effects in a pathway such as the one below.



Stochastic Focusing

Two step cascade reactions [Paulsson, et al. PNAS 97, 7148 (2000)]



$$\text{Sensitivity} = \frac{d\langle v_3\rangle}{d\langle S_1\rangle}\frac{\langle S_1\rangle}{\langle v_3\rangle} = \frac{d\ln\langle v_3\rangle}{d\ln\langle S_1\rangle}.$$

$$Mean(v_3) = p_4 Mean(S_2).$$

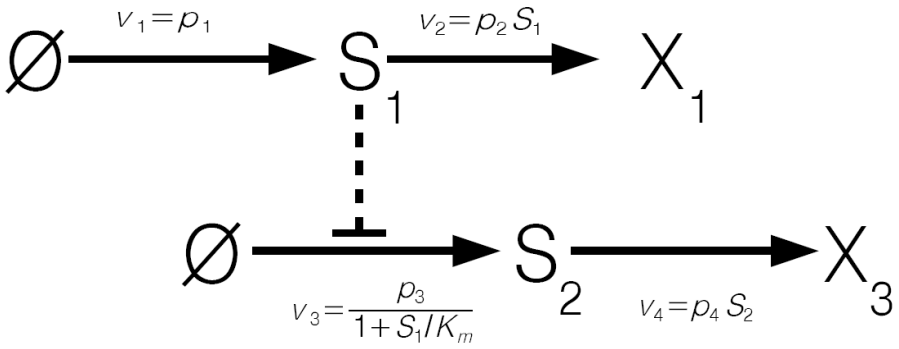
$$Mean(S_2) = \frac{Mean(v_3)}{p_4}.$$

$$\ln[Mean(S_2)] = \ln[Mean(v_3)] - \ln[p_4].$$

$$d\ln[Mean(S_2)] = d\ln[Mean(v_3)] - d\ln[p_4].$$

Stochastic Focusing

Two step cascade reactions [Paulsson, et al. PNAS 97, 7148 (2000)]



$$\text{Sensitivity} = \frac{d\langle S_2\rangle}{d\langle S_1\rangle}\frac{\langle S_1\rangle}{\langle S_2\rangle} = \frac{d\ln\langle S_2\rangle}{d\ln\langle S_1\rangle}.$$