# 2010 q-bio Summer School, Theme 2: Stochastic Biochemistry 

Brian Munsky<br>Center for NonLinear Studies,<br>Information Sciences Group (CCS-3),<br>and the National Flow Cytometry Resource,<br>Los Alamos National Laboratory<br>\section*{brian.munsky@gmail.com}

Stochastic Biochemistry: Theme Overview

1. Stochastic Phenomena: origins and consequences.
2. Single Cell Research.

## Origins of Stochasticity: 1) Small molecular copy numbers



- Proteins build cellular structures, pass cellular information and regulate cellular activities. Variable copy numbers (~0-100,000/cell).
- mRNA transfer instructions for creating specific proteins. Low copy numbers (~0-100/cell).
- DNA contains all of the genetic instructions. Extremely low copy numbers (~0-5/cell).


## The Central Dogma of Molecular Biology

## Origins of Stochasticity:

 2) Spatial fluctuations of cellular constituents.

Thermal fluctuations will lead to randomness in times between reactions.

Origins of Stochasticity:
3) Competition of different events.


## Different reactions will lead to different consequences.



Which ever molecule wins the race will define the reaction.

## Origins of Stochasticity: 4) Extrinsic fluctuations.

Changes in temperature, nutrients, radiation, chemicals, pressure, etc...

Fluctuations of upstream genes, intercellular signals.

## Intrinsic versus Extrinsic Noise

- Variability is present and can be measured


Low Intrinsic Noise

High Intrinsic Noise


Elowitz et al, "Stochastic Gene Expression in a Single Cell", Science 2002

- Inserted two reporter genes on the chromosome (cfp, yfp)
- Each was controlled by the same promoter
- Expression of cfp shown in green, yfp in red


## Stochastic Effects Lead to Phenotypical Differences



## Stochastic Phenomena: <br> 1) Signal Amplification (or damping).



Johan Paulsson , Otto G. Berg, and Måns Ehrenberg, "Stochastic Focusing: Fluctuationenhansed sensitivity of intracellular regulation" PNAS 2000

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


## Stochastic Phenomena: 2) Noise Induced Oscillations

## Circadian rhythm




- 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]


## Stochastic Phenomena:

3) Stochastic Switching
$\rightarrow 0000$
Same chemical environment. $-7000$ Same genetic code.
$\infty \infty$


Random reactions can lead to vastly different results!

Harmless
phenotype.

Highly infectious phenotype.

## The Importance of Single Cell Analyses

## For these systems, we need single cell analyses to answer:

$\star$ What will happen?
$\star$ How frequently?
ڤ Why does it happen?
Ł Under what conditions?
$\star$ What advantages does it provide?
ฝ How can we prevent it?
$\star$ How can we cause it?


Stochastic Biochemistry: Theme Overview

1. Stochastic Phenomena: origins and consequences.
2. Single Cell Research.







Population of mRNA's





Light sensing Bacteria, Voigt Lab, 2005

Native circuit





Stochastic Biochemistry: Lecture Plan

1) Theoretical Techniques
(Munsky, Nemenman, Zilman)
2) Experimental Techniques
(Marrone, Raj, Werner, Voigt)

## Lecture Plan: 1) Theoretical Techniques

- Today and Wednesday--Brian Munsky (LANL-CNLS)
- Modeling of stochastic effects in systems biology.
- Friday, August 6--Ilya Nemenman (Emory)
- Signal processing in biochemical networks: Fourier transforms, central limit theorem, linear feedback, and all that.
- Monday, August 9-- Anton Zilman (LANL-CNLS)
- History of Stochastic Modeling in Physics.
- Advanced stochastic analyses: Fokker Planck equation, Moment Generating Functions, etc...


## Lecture Plan: <br> 2) Experimental Techniques

- Tuesday, August 3--Arjun Raj (U-Penn)
- Measuring cell-to-cell variability with fluorescence microscopy and single molecule Fluorescence In Situ Hybridization (FISH) techniques.
- Tuesday, August 3--Babetta Marrone (LANL-B9)
- Measuring cell-to-cell variability with flow cytometry and fluorescence activated cell sorting.
- Wednesday, August 4--Jim Werner (LANL-CINT)
- Fluorescence Correlation Spectroscopy (FCS) and 3 Dimensional SingleMolecule Tracking
- Wednesday, August 4--Brian Munsky (LANL-CNLS)
- Integrating single cell data and stochastic models.
- Tuesday, August 10--Christopher Voigt (UCSF)
- Synthetic Biology

Lecture 1: Modeling of stochastic gene regulation (Part 1).

## On the menu...

- Today (Part 1)
- Solutions for Simple Stochastic Processes (Transcription)
- Importance of Population Size
- Stochastic Chemical Kinetics
- Moment Computations for Linear Propensities
- Moment Closures for Non-Linear Propensities
- Wednesday (8:40-10:25) (Part 2)
- Monte Carlo Simulation Techniques
* Gillespie (SSA), Tau leaping, Chemical Langevin (SDEs), Slow Scale SSA.
- Density Computations with Finite State Projection Techniques
- Switch and Trajectory Analyses
- Examples and software


## The Central Dogma of Molecular Biology



- Proteins assemble to build cellular structures, pass cellular information and regulate cellular activities.
- mRNA transfer instructions for the creation of specific proteins.
- DNA contains all of the genetic instructions.


## The Central Dogma of Molecular Biology



$$
\begin{aligned}
& \text { Deterministic model } \\
& \frac{d[m R N A]}{d t}=-\gamma_{r}[m R N A]+k_{r} \\
& \frac{d[p r o t e i n]}{d t}=-\gamma_{p}[\text { protein }]+k_{p}[m R N A]
\end{aligned}
$$

## Stochastic model

- Probability a single mRNA is transcribed in time $d t$ is $k_{r} d t$.
- Probability a single mRNA is degraded in time $d t$ is $(\# m R N A) \cdot \gamma_{r} d t$


## Intrinsic Variability in Gene Expression



## 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
Source of variability at cellular level....
- Small \# of molecules
- Random events
"Intrinsic noise"

The Markov Description of Biochemical Processes

A Jump-Markov description of chemical kinetics

- At any time, the state of the system is defined by its integer population vector: $\mathrm{x} \in \mathbb{Z}^{N}$
- Reactions are transitions from one state to another:


A Jump-Markov description of chemical kinetics

- At any time, the state of the system is defined by its integer population vector: $\mathrm{x} \in \mathbb{Z}^{N}$
- Reactions are transitions from one state to another:
- These reactions are random, others could have occurred:


A Jump-Markov description of chemical kinetics


A Jump-Markov description of chemical kinetics


A Jump-Markov description of chemical kinetics


## A Jump-Markov description of chemical kinetics



## Reaction Stoichiometry

- The Stoichiometric vector, s, refers to the relative change in the population vector after a reaction.
- There may be many different reactions for a given stoichiometry.



## Reaction Propensities

- The propensity, $\mathbf{w}$, of a reaction is its rate.
- $\mathbf{w}_{\mu} d t$ is the probability that the $\mu^{t h}$ reaction will occur in a time step of length $d t$.
- Typically, propensities depend only upon reactant populations.

| $\mathbf{s}_{2}=[-1,0]^{T}$ | $w_{2}\left(x_{1}, x_{2}\right)$ |
| :---: | :---: |
| $\mathcal{S}_{1}+\mathcal{S}_{1} \rightarrow \mathcal{S}_{1}$ | $k_{1} x_{2}\left(x_{1}-1\right) / 2$ |
| $\mathcal{S}_{1}+\mathcal{S}_{2} \rightarrow \mathcal{S}_{2}$ | $k_{2} x_{1} x_{2}$ |
| $\mathcal{S}_{1} \rightarrow \emptyset$ | $k_{3} x_{1}$ |



Markov is a forgetful process

## Markov Reaction Times

Probability reaction will occur in $[t, t+\Delta t): \quad w \Delta t+\mathcal{O}(\Delta t)^{2}$
Probability reaction will not occur in $[t, t+\Delta t) \quad 1-w \Delta t+\mathcal{O}(\Delta t)^{2}$
Probability a reaction will not occur in two such time intervals $[t, t+2 \Delta t):\left(1-w \Delta t+\mathcal{O}(\Delta t)^{2}\right)^{2}=1-2 w \Delta t+\mathcal{O}(\Delta t)^{2}$ Suppose that $\tau=K \Delta t$, then the probability that no reaction will occur in the interval $[t, t+\tau)$ is

$$
\left(1-w \frac{\tau}{K}+\mathcal{O}\left(K^{-2}\right)\right)^{K}
$$

Taking the limit as K goes to infinity yields that the probability that no reaction will occur in the interval $[t, t+\tau)$ is

$$
\lim _{k \rightarrow \infty}\left(1-w \frac{\tau}{K}+\mathcal{O}\left(K^{-2}\right)\right)^{K}=\exp (-w \tau)
$$

## Markov Reaction Times

The probability that a reaction will occur in the interval $[t, t+\tau)$ is $F_{T}(\tau)=1-\exp (-w \tau)$. This is a cumulative distribution.

The density (derivative) of the random number, $T$, is:

$$
f_{T}(\tau)=\frac{1}{w} \exp (-w \tau)
$$

Such a random number is known as an exponentially distributed random number.

Notation: $\quad T \in \operatorname{EXP}(\lambda) \rightarrow \quad T$ is an exponentially distributed r.v. with parameter: $\lambda$.

## Markov Reaction Times

- We have assumed that the system is fully described by the population vectors.
- If no reaction occurs, then nothing will have changed.
- Waiting times must be memoryless random variables.

- No matter where we cut and scale the distribution, it must always looks the same.

The exponential is the only continuous r.v. with this property.

## Generating Reaction Times

- To generate an exponentially distributed random number, all we need is a uniform random number generator.
- Find the cumulative distribution,

$$
F(t)=1-\exp (-\lambda t)
$$

- Generate uniform random number,

$$
r \in \mathrm{U}[0,1]
$$

- Find intersection where $F(t)=r$ :

$$
\tau=\frac{1}{\lambda} \log \frac{1}{1-r}
$$

- This is the time of the next reaction.

The (Chemical) Master Equation
(Forward Kolmorogrov Equation)

## The Chemical Master Equation

Prob. that no reactions fire in $[t, t+d t]=1-\sum_{k} w_{k}(x) d t+\mathcal{O}\left(d t^{2}\right)$
Prob. that reaction $R_{k}$ fires once in $[t, t+d t]=w_{k}(x) d t+\mathcal{O}\left(d t^{2}\right)$
Prob. that more than one reaction fires in $[t, t+d t]=\mathcal{O}\left(d t^{2}\right)$

$$
\begin{aligned}
& p(x, t+d t)=\begin{array}{cc}
\text { at } x & \text { No reaction fires } \\
p(x, t)
\end{array}\left(1-\sum_{k} w_{k}(x) d t+\mathcal{O}\left(d t^{2}\right)\right) \\
&+ \sum_{k} \begin{array}{l}
p\left(x-s_{k}, t\right) \\
R_{k} \text { reaction } \\
\text { away from } x
\end{array}\left(\sum_{k} w_{k}(x) d t+\mathcal{O}\left(d t^{2}\right)\right)+\begin{array}{c}
\mathcal{O}\left(d t^{2}\right) \\
\text { more than one }
\end{array} \\
& \text { reaction in } d t
\end{aligned}
$$

$$
p(x, t+d t)-p(x, t)=-p(x, t) \sum_{k} w_{k}(x) d t+\sum_{k} p\left(x-s_{k}, t\right) w_{k}(x) d t+\mathcal{O}\left(d t^{2}\right)
$$

## The Chemical Master Equation

$$
\frac{d p(x, t)}{d t}=-p(x, t) \sum_{k} w_{k}(x)+\sum_{k} p\left(x-s_{k}, t\right) w_{k}\left(x-s_{k}\right)
$$

Example: Transcription and degradation of mRNA

## RNA Copy Number as a Random Variable



Slide Contributed by Mustafa Khammash


Find $p(n, t)$, the probability that $N(t)=n$.

$$
\begin{aligned}
& P(n, t+d t)=P(n-1, t) \cdot k d t \text { Prob. }\{N(t)=n-1 \text { and mRNA created in }[\mathrm{t}, \mathrm{t}+\mathrm{dt})\} \\
&+P(n+1, t) \cdot(n+1) \gamma d t \quad \text { Prob. }\{N(t)=n+1 \text { and mRNA degraded in }[\mathrm{t}, \mathrm{t}+\mathrm{dt})\} \\
&+P(n, t) \cdot(1-k d t)(1-n \gamma d t) \text { Prob. }\{N(t)=n \text { and } \\
&\text { mRNA not created nor degraded in }[\mathrm{t}, \mathrm{t}+\mathrm{dt})\}
\end{aligned}
$$

$$
P(n, t+d t)-P(n, t)=P(n-1, t) k d t+P(n+1, t)(n+1) \gamma d t-P(n, t)(k+n \gamma) d t
$$

$$
+O\left(d t^{2}\right)
$$

Dividing by $d t$ and taking the limit as $d t \rightarrow 0$

$$
\begin{aligned}
& \text { The Chemical Master Equation } \\
& \qquad \frac{d}{d t} P(n, t)=k P(n-1, t)+(n+1) \gamma P(n+1, t)-(k+n \gamma) P(n, t)
\end{aligned}
$$

## mRNA Stationary Distribution

We look for the stationary distribution $\quad P(n, t)=p(n) \forall t$
The stationary solution satisfies: $\frac{d}{d t} P(n, t)=0$
From the Master Equation ...

$$
\begin{array}{cc} 
& (k+n \gamma) p(n)=k p(n-1)+(n+1) \gamma p(n+1) \\
n=0 & k p(0)=\gamma p(1) \\
n=1 & k p(1)=2 \gamma p(2) \\
n=2 & k p(2)=3 \gamma p(3) \\
\vdots & \\
& k p(n-1)=n \gamma p(n)
\end{array}
$$

$k p(n-1)=n \gamma p(n)$ We can express $p(n)$ as a function of $p(0)$ :

$$
\begin{aligned}
p(n) & =\frac{k}{\gamma} \frac{1}{n} p(n-1) \\
& =\left(\frac{k}{\gamma}\right)^{2} \frac{1}{n} \frac{1}{n-1} p(n-2) \\
& \vdots \\
& =\left(\frac{k}{\gamma}\right)^{n} \frac{1}{n!} p(0)
\end{aligned}
$$

We can solve for $p(0)$ using the fact $\sum_{n=0}^{\infty} p(n)=1$

$$
\begin{aligned}
1 & =\sum_{n=0}^{\infty}\left(\frac{k}{\gamma}\right)^{n} \frac{1}{n!} p(0) \quad n=0 \\
& =e^{k / \gamma} p(0) \quad \Rightarrow(0)=e^{-k / \gamma} \\
p(n) & =e^{-a} \frac{a^{n}}{n!} \quad a=\frac{k}{\gamma} \quad \quad \text { Poisson Distribution }
\end{aligned}
$$

We can compute the mean and variance of the Poisson RV $\bar{N}$ with density $p(n)=e^{-a \frac{a^{n}}{n!}}$ :

$$
\mu=E[\bar{N}]=\sum_{n=0}^{\infty} n p(n)=e^{-a} \sum_{n=0}^{\infty} n \frac{a^{n}}{n!}=a
$$

The second moment

$$
E\left[\bar{N}^{2}\right]=\sum_{n=0}^{\infty} n^{2} p(n)=a^{2}+a
$$

Therefore,

$$
\begin{gathered}
\sigma^{2}=E\left[\bar{N}^{2}\right]-E[\bar{N}]^{2}=a \\
\text { mean }=\text { variance }=a
\end{gathered}
$$

The coefficient of variation $C_{v}=\sigma / \mu$ is

$$
C_{v}=\frac{1}{\sqrt{a}}=\frac{1}{\sqrt{\mu}}
$$



The Relationship of Deterministic to Stochastic Biochemical Processes.

## Relationship of Stochastic and Deterministic Descriptions

Given $N$ species $X_{1}, \ldots, X_{N}$ and $M$ elementary reactions. Let $\Phi_{i}:=\left[X_{i}\right]$.

A deterministic description can be obtained from mass-action kinetics:

$$
\frac{d \Phi}{d t}=S f(\Phi)
$$

where $f(\cdot)$ is at most a second order monomial. It depends on the type of reactions and their rates.

## Example:

$$
\begin{array}{r}
A+B \xrightarrow{k_{1}} C \\
A \xrightarrow{k_{2}} B
\end{array}
$$

$\frac{d \Phi_{A}}{d t}=-k_{1} \Phi_{A} \Phi_{B}-k_{2} \Phi_{A}$

$$
\frac{d \Phi}{d t}=S f(\Phi) \text { where }
$$

$\frac{d \Phi}{d t} \mathrm{~B}=-k_{1} \Phi_{A} \Phi_{B}+k_{2} \Phi_{A}$
or $\quad S=\left[\begin{array}{cc}-1 & -1 \\ -1 & 1 \\ 1 & 0\end{array}\right], f(\Phi)=\left[\begin{array}{c}k_{1} \Phi_{A} \Phi_{B} \\ k_{2} \Phi_{A}\end{array}\right]$

## Relationship of Stochastic and <br> Deterministic Descriptions

Define $X^{\Omega}(t)=\frac{X(t)}{\Omega}$.
Question: How does $X^{\Omega}(t)$ relate to $\Phi(t)$ ?

Fact: Let $\Phi(t)$ be the deterministic solution to the reaction rate equations

$$
\frac{d \Phi}{d t}=S f(\Phi), \Phi(0)=\Phi_{0} .
$$

Let $X^{\Omega}(t)$ be the stochastic representation of the same chemical systems with $X^{\Omega}(0)=\Phi_{0}$. Then for every $t \geq 0$ :

$$
\lim _{\Omega \rightarrow \infty} \sup _{s \leq t}\left|X^{\Omega}(s)-\Phi(s)\right|=0 \text { a.s. }
$$

$x$ produced with rate $k(x)$ and degraded with rate $\gamma_{0} x$.

$w_{1}(\phi)=\gamma \phi$
$w_{2}(\phi)=\left(20+40 \frac{\phi^{10}}{40^{10}+\phi^{10}}\right)$
Deterministic


$$
\begin{aligned}
& w_{1}(X)=\Omega \gamma_{0} X / \Omega=\gamma_{0} X \\
& w_{2}(X)=\Omega\left(20+40 \frac{(X / \Omega)^{10}}{40^{10}+(X / \Omega)^{10}}\right)
\end{aligned}
$$

Stochastic

## Moment Computations

- Affine Propensity
- Moment Closures


## Moment Computations

For the first moment $E\left[X_{i}\right]$, multiply the CME by $x_{i}$ and sum over all $\left(x_{1}, \ldots, x_{N}\right) \in \mathbb{N}^{N}$

For the second moment $E\left[X_{i} X_{j}\right]$, multiply the CME by $x_{i} x_{j}$ and sum over all $\left(x_{1}, \ldots, x_{N}\right) \in \mathbb{N}^{N}$

$$
\begin{aligned}
& \frac{d E\left[X_{i}\right]}{d t}=\sum_{k=1}^{M} s_{i k} E\left[w_{k}(X)\right] \\
& \frac{d E\left[X_{i} X_{j}\right]}{d t}=\sum_{k=1}^{M}\left(s_{i k} E\left[X_{j} w_{k}(X)\right]+E\left[X_{i} w_{k}(X)\right] s_{j k}+s_{i k} s_{j k} E\left[w_{k}(X)\right]\right) \\
& \text { Let } w(x)=\left[w_{1}(x), \ldots, w_{M}(x)\right]^{T}
\end{aligned}
$$

In matrix notation:

$$
\begin{aligned}
\frac{d E[X]}{d t} & =S E[w(X)] \\
\frac{d E\left[X X^{T}\right]}{d t} & =S E\left[w(X) X^{T}\right]+E\left[w(X) X^{T}\right]^{T} S^{T}+S\{\operatorname{diag} E[w(X)]\} S^{T}
\end{aligned}
$$

## Affine Propensity

Suppose the propensity function is affine:

$$
w(x)=W x+w_{0}, \quad\left(W \text { is } N \times N, w_{0} \text { is } N \times 1\right)
$$

Then $E[w(X)]=W E[X]+w_{0}$, and $E\left[w(X) X^{T}\right]=W E\left[X X^{T}\right]+w_{0} E\left[X^{T}\right]$.
This gives us the moment equations:

$$
\begin{array}{rlr}
\frac{d}{d t} E[X] & =S W E[X]+S w_{0} & \text { First Moment } \\
\frac{d}{d t} E\left[X X^{T}\right] & =S W E\left[X X^{T}\right]+E\left[X X^{T}\right] W^{T} S^{T}+S \operatorname{diag}\left(W E[X]+w_{0}\right) S^{T} \\
& +S w_{0} E\left[X^{T}\right]+E[X] w_{0}^{T} S^{T} & \text { Second Moment }
\end{array}
$$

These are linear ordinary differential equations and can be easily solved!

## Affine Propensity (cont.)

Define the covariance matrix $\Sigma=E\left[(X-E[X])(X-E(X)]^{T}\right]$.
We can also compute covariance equations:

$$
\frac{d}{d t} \Sigma=S W \Sigma+\Sigma W^{T} S^{T}+S \operatorname{diag}\left(W E[X]+w_{0}\right) S^{T}
$$

## Steady-state Case

The steady-state moments and covariances can be obtained by solving linear algebraic equations:

Let $\bar{X}=\lim _{t \rightarrow \infty} E[X(t)]$ and $\bar{\Sigma}=\lim _{t \rightarrow \infty} \Sigma(t)$.
Then

$$
\begin{gathered}
S W \bar{X}=-S w_{0} \\
S W \bar{\Sigma}+\bar{\Sigma} W^{T} S^{T}+S \operatorname{diag}\left(W \bar{X}+w_{0}\right) S^{T}=0
\end{gathered}
$$

## Fluctuations Arise from Noise Driven Dynamics

Define $A=S W$, and $B=S \sqrt{\operatorname{diag}\left(W \bar{X}+w_{0}\right)}$.
The steady-state covariances equation

$$
S W \bar{\Sigma}+\bar{\Sigma} W^{T} S^{T}+S \operatorname{diag}\left(W \bar{X}+w_{0}\right) S^{T}=0
$$

becomes

$$
A \bar{\Sigma}+\bar{\Sigma} A^{T}+B B^{T}=0 \quad \text { Lyapunov Equation }
$$

## Example: Gene Expression

## Application to Gene Expression



## Steady-State Moments

$$
A=S W=\left[\begin{array}{cc}
-\gamma_{r} & 0 \\
k_{p} & -\gamma_{p}
\end{array}\right], \quad S w_{0}=\left[\begin{array}{c}
k_{r} \\
0
\end{array}\right]
$$

$\bar{X}=-A^{-1} S w_{0}=\left[\begin{array}{c}\frac{k_{r}}{\gamma_{r}} \\ \\ \frac{k_{p} k_{r}}{\gamma_{p} \gamma_{r}}\end{array}\right]$

## Steady-State Covariance

$B B^{T}=S \operatorname{diag}\left(W \bar{X}+w_{0}\right) S^{T}=\left[\begin{array}{cc}2 k_{r} & 0 \\ 0 & \frac{2 k_{p} k_{r}}{\gamma_{r}}\end{array}\right]$
The steady-state covariances equation

$$
A \bar{\Sigma}+\bar{\Sigma} A^{T}+B B^{T}=0 \quad \text { Lyapunov Equation }
$$

can be solved algebraically for $\bar{\Sigma}$.

$$
\bar{\Sigma}=\left[\begin{array}{cc}
\frac{k_{r}}{\gamma_{r}} & \frac{k_{p} k_{r}}{\gamma_{r}\left(\gamma_{r}+\gamma_{p}\right)} \\
\frac{k_{p} k_{r}}{\gamma_{r}\left(\gamma_{r}+\gamma_{p}\right)} & \frac{k_{p} k_{r}}{\gamma_{p} \gamma_{r}}\left(1+\frac{k_{p}}{\gamma_{r}+\gamma_{p}}\right)
\end{array}\right]
$$

## Coefficients of Variation

$$
\begin{aligned}
& C_{v r}^{2}=\frac{1}{\frac{k_{r}}{\gamma_{r}}}=\frac{1}{\bar{X}_{1}} \\
& C_{v p}^{2}=\frac{1}{\frac{k_{r} k_{p}}{\gamma_{r} \gamma_{p}}}\left(1+\frac{k_{p}}{\gamma_{r}+\gamma_{p}}\right)=\frac{1}{\bar{X}_{2}}\left(1+\frac{k_{p}}{\gamma_{r}+\gamma_{p}}\right)
\end{aligned}
$$

Question: Does a large $\bar{X}_{2}$ imply a small $C_{v p}$ ?

$$
\begin{aligned}
C_{v p}^{2} & =\frac{1}{\frac{k_{r} k_{p}}{\gamma_{r} \gamma_{p}}}\left(1+\frac{k_{p}}{\gamma_{r}+\gamma_{p}}\right) \\
& \geq \frac{1}{\frac{k_{r} k_{p}}{\gamma_{r} \gamma_{p}}}\left(\frac{k_{p}}{\gamma_{r}+\gamma_{p}}\right)=\frac{\gamma_{r} \gamma_{p}}{k_{r}} \cdot \frac{1}{\gamma_{r}+\gamma_{p}}
\end{aligned}
$$

$\bar{X}_{2}=\frac{k_{r} k_{p}}{\gamma_{r} \gamma_{p}}$, which can be chosen independently from $C_{v p}$.

Large mean does not imply small fluctuations!

$$
\mathbb{E}\{P\}=100, \quad \gamma_{r}=\gamma_{p}=1
$$




$$
k_{r}=100 \quad k_{p}=1
$$

$$
C_{v p}^{2}=0.015
$$

$$
k_{r}=0.1 \quad k_{p}=1000
$$

$$
\frac{P}{\mathbb{E}\{P\}}{ }_{5}^{15}{ }_{5}^{10}
$$

## Moment Computations

- Affine Propensity
- Moment Closures


## Moment Closures.

From before, the mean level changes as:
$\frac{d E[X]}{d t}=S E[w(X)]$

- When Second and Higher order terms exist in the propensity functions, each moment depends upon higher moments.
- For example, if $w(X)=\mathbf{u} X^{T} X \mathbf{v}$, then
$\frac{d E[X]}{d t}=S \mathbf{u} E\left[X^{T} X\right] \mathbf{v}$
- The first moment depends upon the second; the second upon the third; and so on.
- Moment closures are approximations that attempt to remove this dependence.


## Moment Closures.

$$
\begin{aligned}
& \frac{d E\left[X_{j}\right]}{d t}=\sum_{k=1}^{M} s_{k} E\left[w_{k}(X)\right] \\
& \frac{d E\left[X_{i} X_{j}\right]}{d t}=\sum_{k=1}^{M}\left(s_{i k} E\left[X_{j} w_{k}(X)\right]+E\left[X_{i} w_{k}(X)\right]_{j k}+s_{i k} s_{j k} E\left[w_{k}(X)\right]\right) \\
& \frac{d}{d t}\left[\begin{array}{l}
\left\{\mu_{i}\right\} \\
\left\{\sigma_{i j}\right\}
\end{array}\right]=\left[\begin{array}{l}
f_{1}\left(\left\{\mu_{i}\right\},\left\{\sigma_{i j}\right\}\right)+u_{1}\left(\left\{\sigma_{i j k}\right\},\left\{\sigma_{i j k l}\right\}, \ldots\right) \\
f_{2}\left(\left\{\mu_{i}\right\},\left\{\sigma_{i j}\right\}\right)+u_{2}\left(\left\{\sigma_{i j k}\right\},\left\{\sigma_{i j k l}\right\}, \ldots\right)
\end{array}\right], \\
& \frac{d}{d t}\left[\begin{array}{c}
\left\{\mu_{i}\right\} \\
\left\{\sigma_{i j}\right\}
\end{array}\right]=\left[\begin{array}{l}
f_{1}\left(\left\{\mu_{i}\right\},\left\{\sigma_{i j}\right\}\right)+\widehat{u}_{1}\left(\left\{\mu_{i}\right\},\left\{\sigma_{i j}\right\}\right) \\
f_{2}\left(\left\{\mu_{i}\right\},\left\{\sigma_{i j}\right\}\right)+\widehat{u}_{2}\left(\left\{\mu_{i}\right\},\left\{\sigma_{i j}\right\}\right)
\end{array}\right],
\end{aligned}
$$

where the choice of $\widehat{u}_{1}$ and $\widehat{u}_{2}$
depends upon the chosen moment closure.

## Gaussian Moment Closure

- If one assumes that the distributions are Gaussian, then the closure is simple:

$$
\sigma_{i j k}=\mathbb{E}\left\{\left(X_{i}-\mathbb{E}\left\{X_{i}\right\}\right)\left(X_{j}-\mathbb{E}\left\{X_{j}\right\}\right)\left(X_{k}-\mathbb{E}\left\{X_{k}\right\}\right)\right\}=0
$$

- which yields:

$$
\begin{array}{r}
\mathbb{E}\left\{\left(X_{i} X_{j} X_{k}\right\}=-\mathbb{E}\left\{X_{i} X_{j}\right\} \mathbb{E}\left\{X_{k}\right\}-\mathbb{E}\left\{X_{j} X_{k}\right\} \mathbb{E}\left\{X_{i}\right\}\right. \\
-\mathbb{E}\left\{X_{k} X_{i}\right\} \mathbb{E}\left\{X_{j}\right\}+2 \mathbb{E}\left\{X_{i}\right\} \mathbb{E}\left\{X_{j}\right\} \mathbb{E}\left\{X_{k}\right\}
\end{array}
$$

- Higher moments are easy to derive with a moment generating function:

$$
\begin{gathered}
M_{\mathbf{x}}(\mathbf{t})=\exp \left(\mu^{T} \mathbf{t}+1 / 2 \mathbf{t}^{T} \boldsymbol{\Sigma} \mathbf{t}\right) \\
\mathbb{E}\left\{x_{1}^{n_{1}} \ldots x_{4}^{n_{4}}\right\}=\left.\frac{d^{n_{1}+\ldots+n_{4}}}{d x_{1}^{n_{1}} \ldots d x_{4}^{n_{4}}} M_{x}(\mathbf{t})\right|_{\mathbf{t}=\mathbf{0}}
\end{gathered}
$$

## Many other closures are possible:

- If one assumes that the distributions are Log-Normal, a different closure is used:

$$
\mathbb{E}\left[X_{i} X_{j} X_{k}\right]=\frac{\mathbb{E}\left[X_{i} X_{j}\right] \mathbb{E}\left[X_{j} X_{k}\right] \mathbb{E}\left[X_{i} X_{k}\right]}{\mathbb{E}\left[X_{i}\right] \mathbb{E}\left[X_{j}\right] \mathbb{E}\left[X_{k}\right]}
$$

- One of the most common closures is the Linear Noise Approximation.
- In this, all moments are written in terms of themselves and lower moments:
- the mean is set equal to the deterministic process.
- the second moments are assumed to be gaussian, and depend upon the mean and itself.

$$
\frac{d}{d t}\left[\begin{array}{c}
\left\{\mu_{i}\right\} \\
\left\{\sigma_{i j}\right\}
\end{array}\right]=\left[\begin{array}{c}
f_{1}\left(\left\{\mu_{i}\right\}\right) \\
f_{2}\left(\left\{\mu_{i},\left\{\sigma_{i j}\right\}\right)\right.
\end{array}\right]
$$

# Noise Suppression and Exploitation (Examples) 

- Feedback for Noise Suppression
- Stochastic Focussing
- Stochastic Switches


## Noise Attenuation through Negative Feedback

Reactants
$X_{1}(t)$ is \# of mRNA; $X_{2}(t)$ is \# of protein

## Reactions

$R_{1}: \phi \xrightarrow{k_{r}} m R N A \quad k_{r}=k_{0}-k_{1} \cdot(\#$ protein $)$
$R_{2}: m R N A \xrightarrow{\gamma_{r}} \phi$
$R_{3}: m R N A \xrightarrow{k_{p}}$ protein $+m R N A$
$R_{4}:$ protein $\xrightarrow{\gamma_{p}} \phi$
Stoichiometry and Propensity

$$
S=\left[\begin{array}{cccc}
1 & -1 & 0 & 0 \\
0 & 0 & 1 & -1
\end{array}\right]
$$

$$
w(X)=\underbrace{\left[\begin{array}{c}
k_{0}-k_{1} X_{2} \\
\gamma_{r} X_{1} \\
k_{p} X_{1} \\
\gamma_{p} X_{2}
\end{array}\right]}_{W}=\underset{w_{0}}{\left[\begin{array}{cc}
0 & -k_{1} \\
\gamma_{r} & 0 \\
k_{p} & 0 \\
0 & \gamma_{p}
\end{array}\right]}\left[\begin{array}{c}
X_{1} \\
X_{2}
\end{array}\right]+\underset{k_{0}}{\left[\begin{array}{c}
k_{0} \\
0 \\
0 \\
0
\end{array}\right]}
$$

## Steady-State Moments

$$
\begin{aligned}
& A=S W=\left[\begin{array}{cc}
-\gamma_{r} & -k_{1} \\
k_{p} & -\gamma_{p}
\end{array}\right], \quad S w_{0}=\left[\begin{array}{c}
k_{0} \\
0
\end{array}\right] \\
& \bar{X}=-A^{-1} S w_{0}=\left[\begin{array}{c}
\frac{k_{0}}{\gamma_{r}} \\
1+\frac{k_{1} k_{p}}{\gamma_{p} \gamma_{r}} \\
\frac{k_{0} k_{p}}{\gamma r \gamma_{p}} \\
1+\frac{k_{1} k_{p}}{\gamma_{p} \gamma_{r}}
\end{array}\right]=:\left[\begin{array}{l}
\mu_{r} \\
\mu_{p}
\end{array}\right] \\
& \text { Steady-State Covariance } \\
& B B^{T}=S \operatorname{diag}\left(W \bar{X}+w_{0}\right) S^{T}=\left[\begin{array}{cc}
k_{0}+\gamma_{r} \mu_{r}-k_{1} \mu_{p} & 0 \\
0 & k_{p} \mu_{r}+\gamma_{p} \mu_{p}
\end{array}\right]
\end{aligned}
$$

The steady-state covariances equation

$$
A \bar{\Sigma}+\bar{\Sigma} A^{T}+B B^{T}=0 \quad \text { Lyapunov Equation }
$$

can be solved algebraically for $\bar{\Sigma}$.

$$
\Sigma_{22}=\sigma_{p}^{2}=\left[\frac{1-\phi}{1+b \phi} \cdot \frac{b}{1+\eta}+1\right] \mu_{p} \quad \text { where } \phi=\frac{k_{1}}{\gamma_{p}}, b=\frac{k_{p}}{\gamma_{r}}, \eta=\frac{\gamma_{p}}{\gamma_{r}}
$$

## Feedback vs. No Feedback

In order to compare the noise in the two cases, we must ensure that both configuations have the same mean!

Impose the constraint: $\mu_{p}^{F B}=\mu_{p}^{N F B}=: \mu_{p}^{*}$
This may be achieved by choosing $k_{0}=k_{r}+k_{1} \mu_{p}^{N F B}$.


Mean
Variance $\quad\left[\frac{b}{1+\eta}+1\right] \mu_{p}^{*}$

$\mu_{p}^{*}$

$$
\left[\frac{1-\phi}{1+b \phi} \cdot \frac{b}{1+\eta}+1\right] \mu_{p}^{*} \quad \text { where } \phi=\frac{k_{1}}{\gamma_{p}}
$$

Protein variance is always smaller with negative feedback!

## Example



Note that these distributions are NOT Gaussian.

## Exploiting the Noise:

## Failure of the linear noise approximation

$$
\begin{aligned}
& \phi \underset{k_{a} S}{\stackrel{k}{\rightleftharpoons}} I \xrightarrow{k_{p}} P \xrightarrow{1} \phi \\
& \phi \stackrel{k_{s}}{\underset{k_{d}}{\rightleftharpoons}} S
\end{aligned}
$$

may be approximated by

$$
\begin{aligned}
& \phi \xrightarrow{k q} P \xrightarrow{1} \phi \\
& q=\frac{1}{1+\frac{n}{\Omega K}} \quad \begin{array}{l}
K=k_{p} / k_{a} \\
\text { n is \#S }
\end{array} \\
& \text { convex }
\end{aligned}
$$

From Jensen's Inequality:

$$
E[q]=E\left[\frac{1}{1+\frac{n}{\Omega K}}\right] \geq \frac{1}{1+\frac{E[n]}{\Omega K}}
$$

- Noise enhances signal!



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

