DNA 17 Tutorial – Caltech – September, 19, 2011

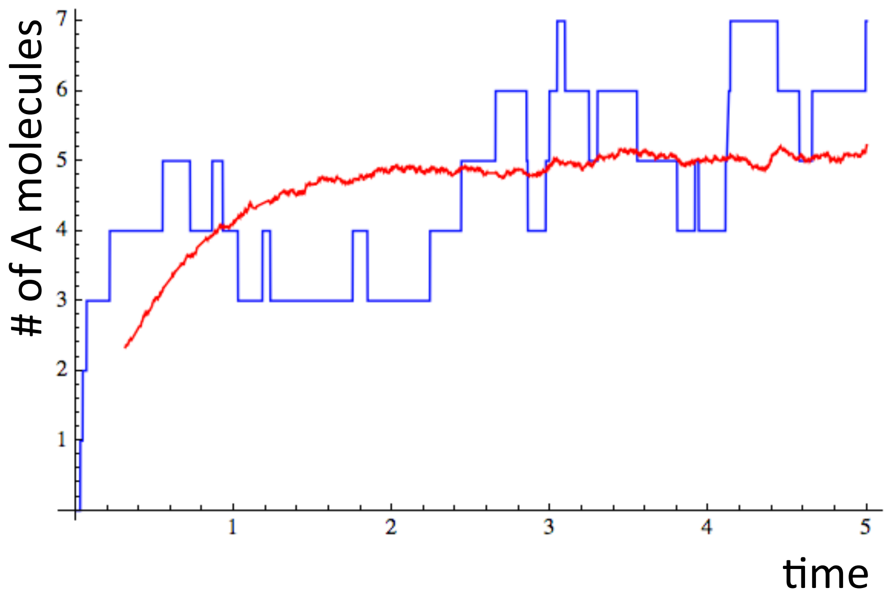
Specification and Control of Stochastic Biochemical Systems

Eric Klavins

University of Washington

Chemical Reactions are Stochastic

$$A \underset{k}{\overset{k}{\rightleftharpoons}} B$$



Single trajectory with 10 molecules

Average of 150 trajectories

At equilibrium:

- There are as many A molecules as B molecules.
- The number of forward reactions balances the number of reverse reactions.

In bulk, with a few nanomoles of molecules in solution, you do not see the fluctuations.

But if had only a few molecules, you would see things differently.

Outline

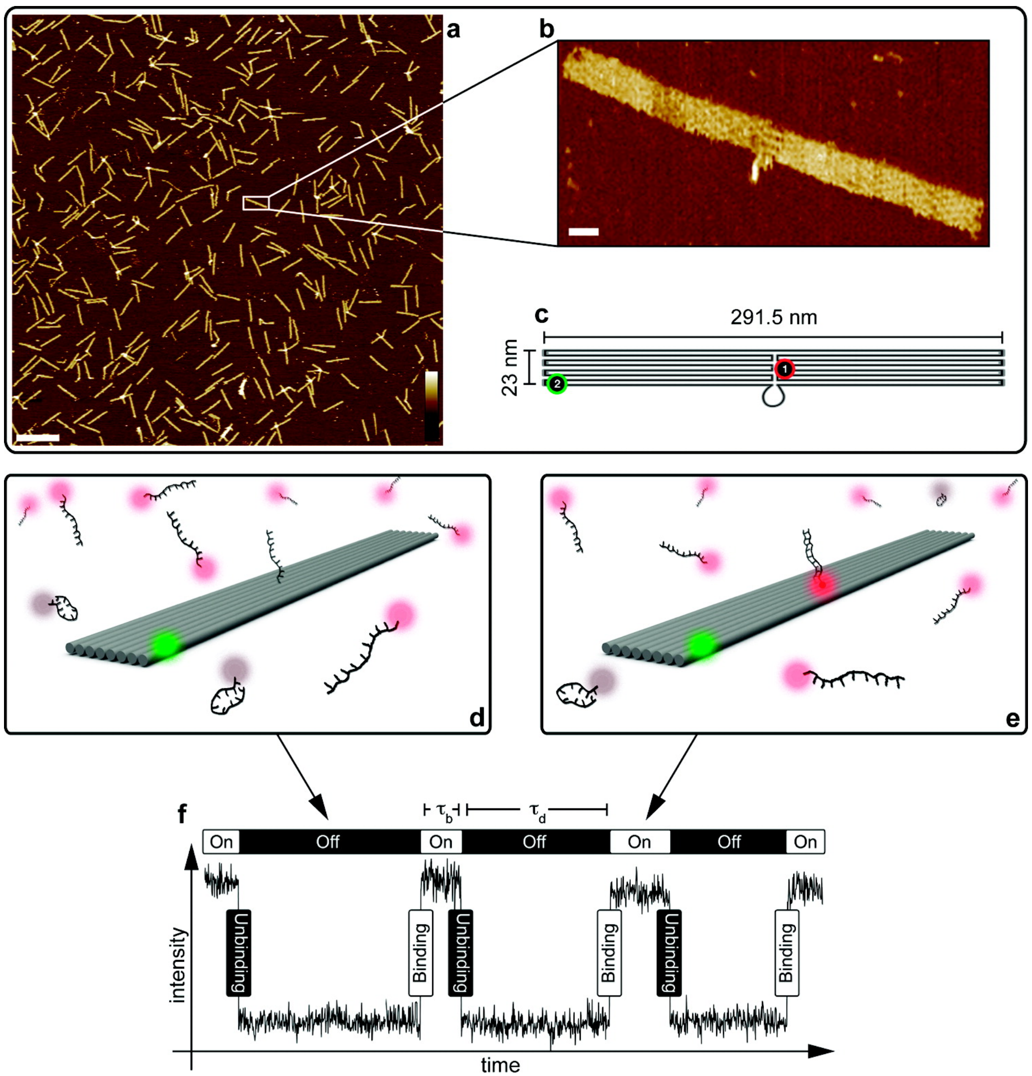
- Example Experimental Systems
- What Stochasticity Can Do
- Analytical Approaches
  - The Master Equation
  - Moment Dynamics
- Simulation Based Approaches
  - Simulation Methods
  - Approximate Abstraction/Refinement

Example:
Arithmetic

Software
Tools

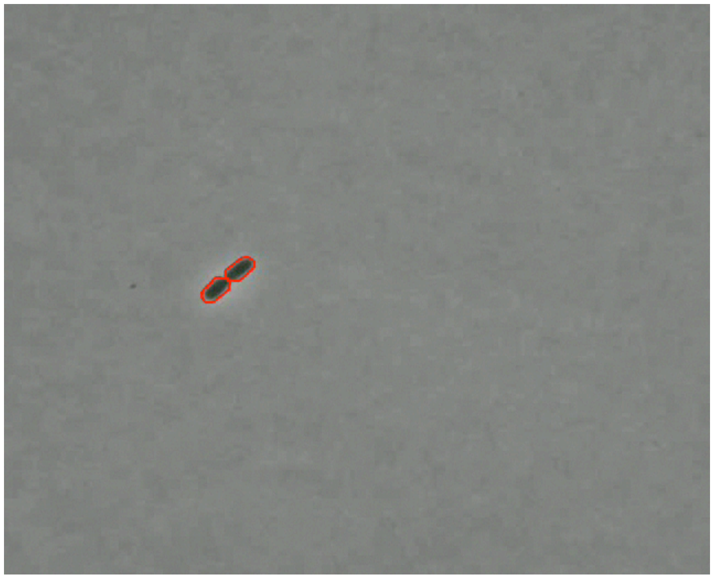
Running Example:
Control of Gene Expression

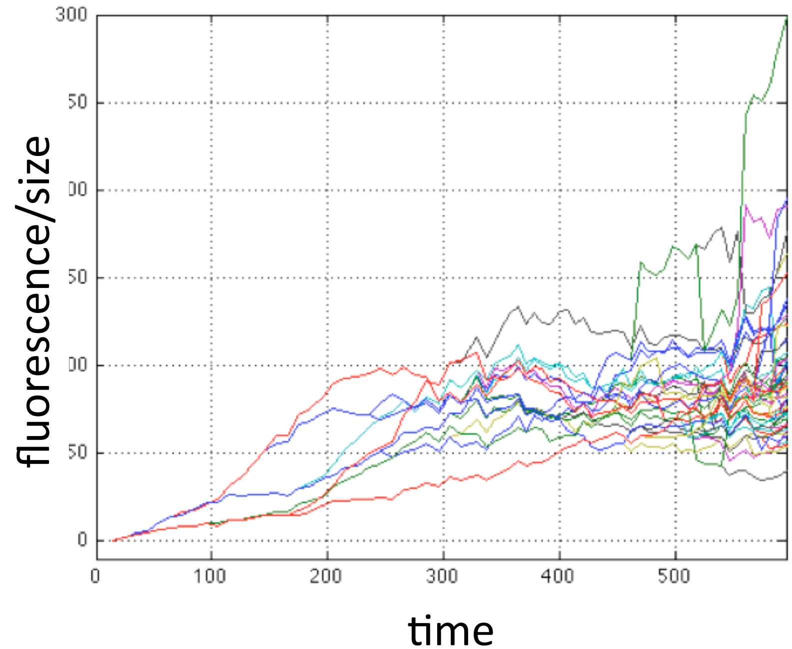
Example: Single Molecule DNA Kinetics Are Observable



R. Jungmann, C. Steinhauer, M. Scheible, A. Kuzyk, P. Tinnefeld, F. C. Simmel, **Single-Molecule Kinetics and Super-Resolution Microscopy by Fluorescence Imaging of Transient Binding on DNA Origami**, *Nano Letters* 10, 4756-4761 (2010)

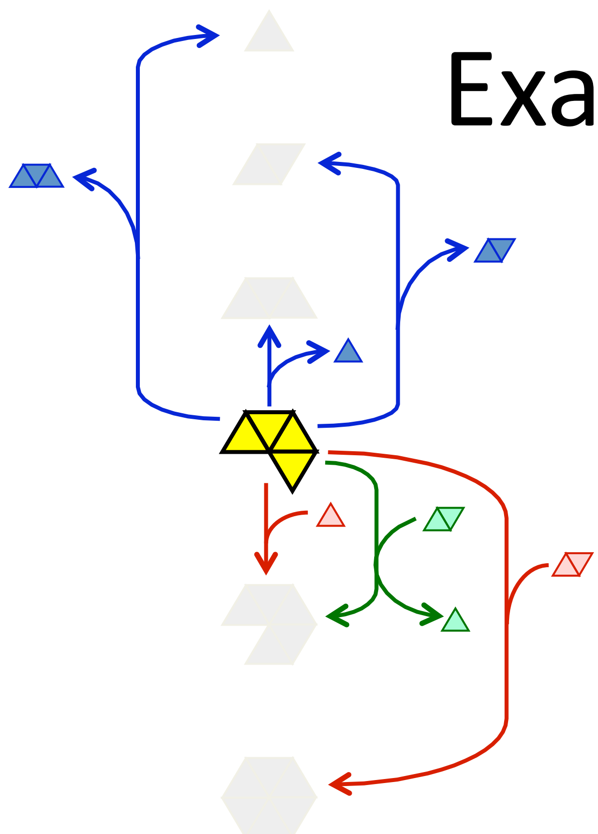
Example: Low Copy Numbers in Cells

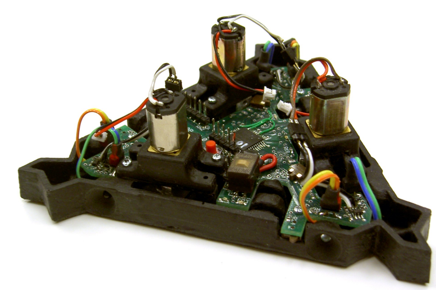


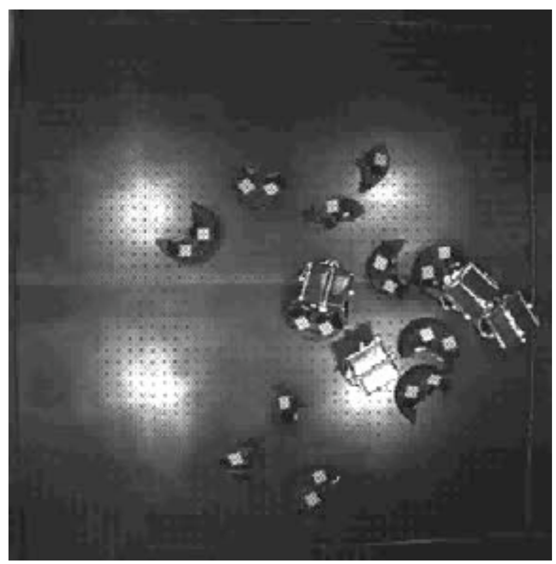


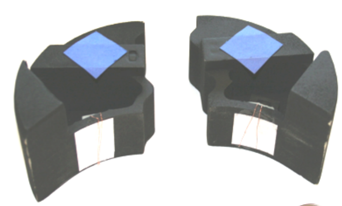
$$
\begin{tikzcd}
\varnothing \arrow[r, "\alpha_R"'] & R \arrow[r, "\alpha_P"'] \arrow[d, "\beta_R"] & P \arrow[r, "\beta_P"'] & \varnothing \\
 & \varnothing & &
\end{tikzcd}
$$

Example: Chemical Robotics









$$P_1 + P_2 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} P_1P_2$$

$$P_1P_2 + A \xrightarrow{k_2} P_1 + P_2 + A'$$

$$A' \xrightarrow{u} A$$

$$u = N_{12}$$

Outline

- Example Experimental Systems
- What Stochasticity Can Do
- Analytical Approaches
    - The Master Equation
    - Moment Dynamics
- Simulation Based Approaches
    - Simulation Methods
    - Approximate Abstraction/Refinement

Running Example:
Control of Gene Expression

Example:
Arithmetic

Software
Tools

What Can You Make?

- Low copy number systems give you integer-valued variables.
  - So you can have states and registers.

- Low copy number systems can flip coins.
  - So you can implement randomized algorithms.

Example:
Computation via Register Machines

states S_0, S_1, ..., S_n

registers R_0, R_1, ..., R_m

$\mathrm{inc}(i, r, j)$

if the state is i, then increase register r and go to state j

$$\boxed{S_i \xrightarrow{k} S_j + M_r}$$

$\mathrm{dec}(i, r, j, k)$

if the state is i and register r is greater than zero, then decrease register r and go to state j; otherwise go to state k

$$\boxed{\begin{array}{l} S_i + M_r \xrightarrow{k} S_j \\ S_i \xrightarrow{\varepsilon} S_k \end{array}}$$

David Soloveichik, Matthew Cook, Erik Winfree and Jehoshua Bruck, **Computation with finite stochastic chemical reaction networks,** *NATURAL COMPUTING.* Volume 7, Number 4, 615-633, 2008.

Example: Multiplication

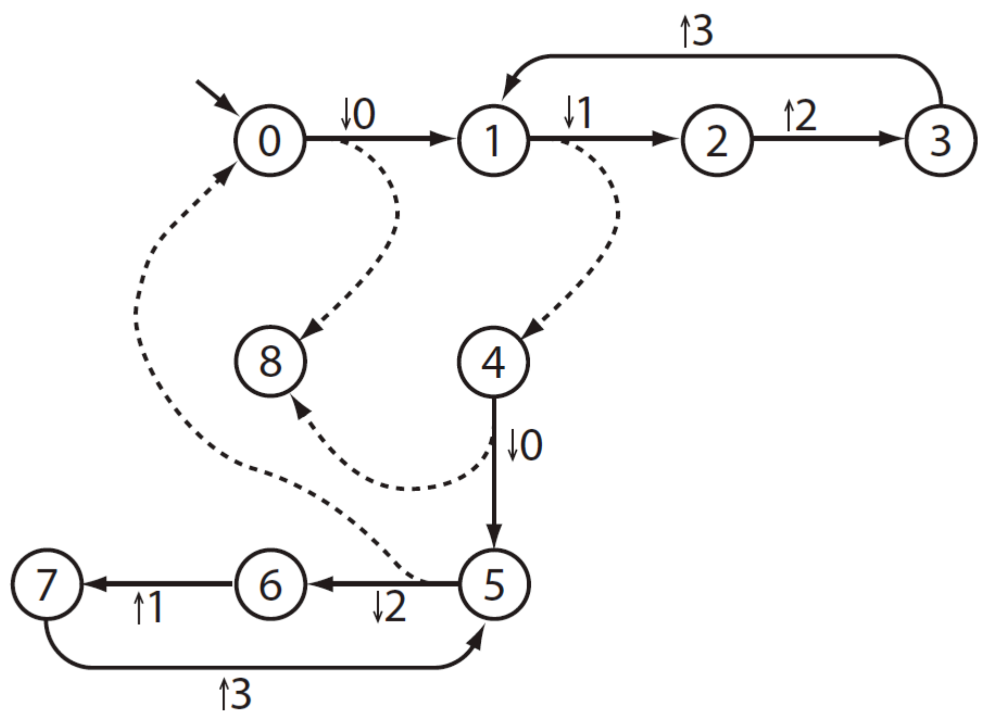
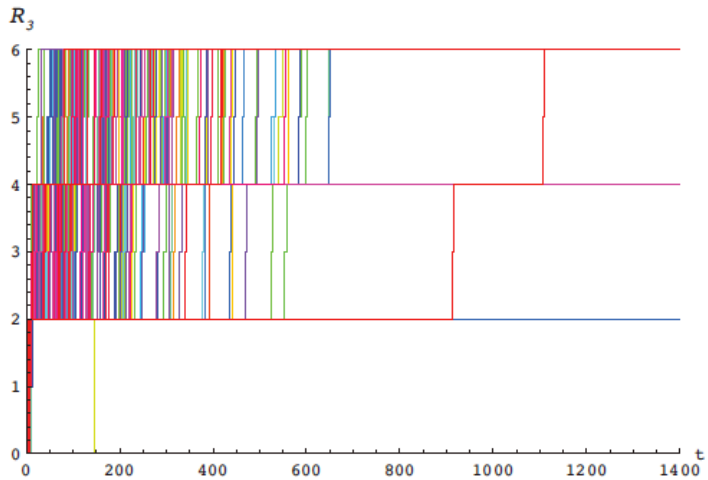


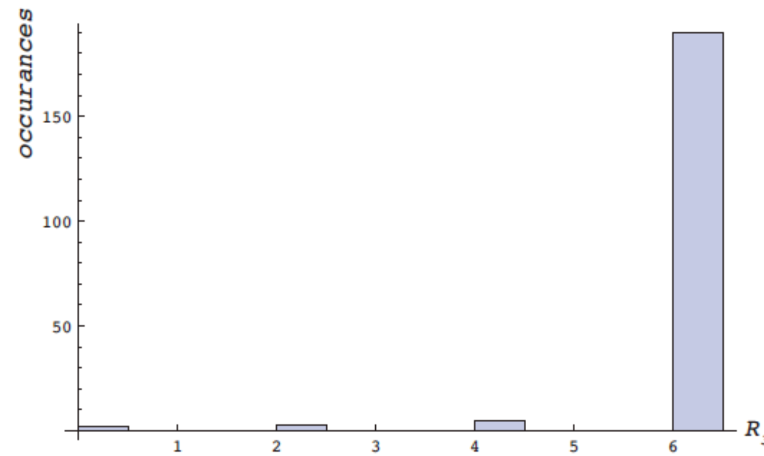
FIGURE 6. A register machine that multiplies the initial contents of registers 0 and 1. Register 3 holds the final value and register 2 is swap space.

Klavins: Course notes.

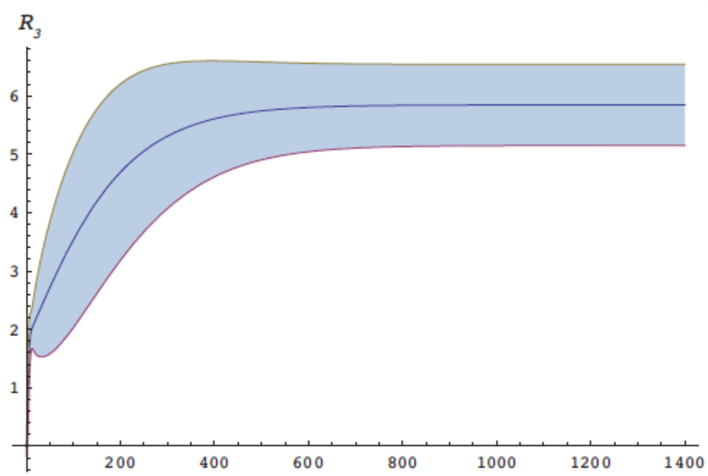
Behavior of the Multiplier



(a)



(b)



(c)

Klavins: Course notes.

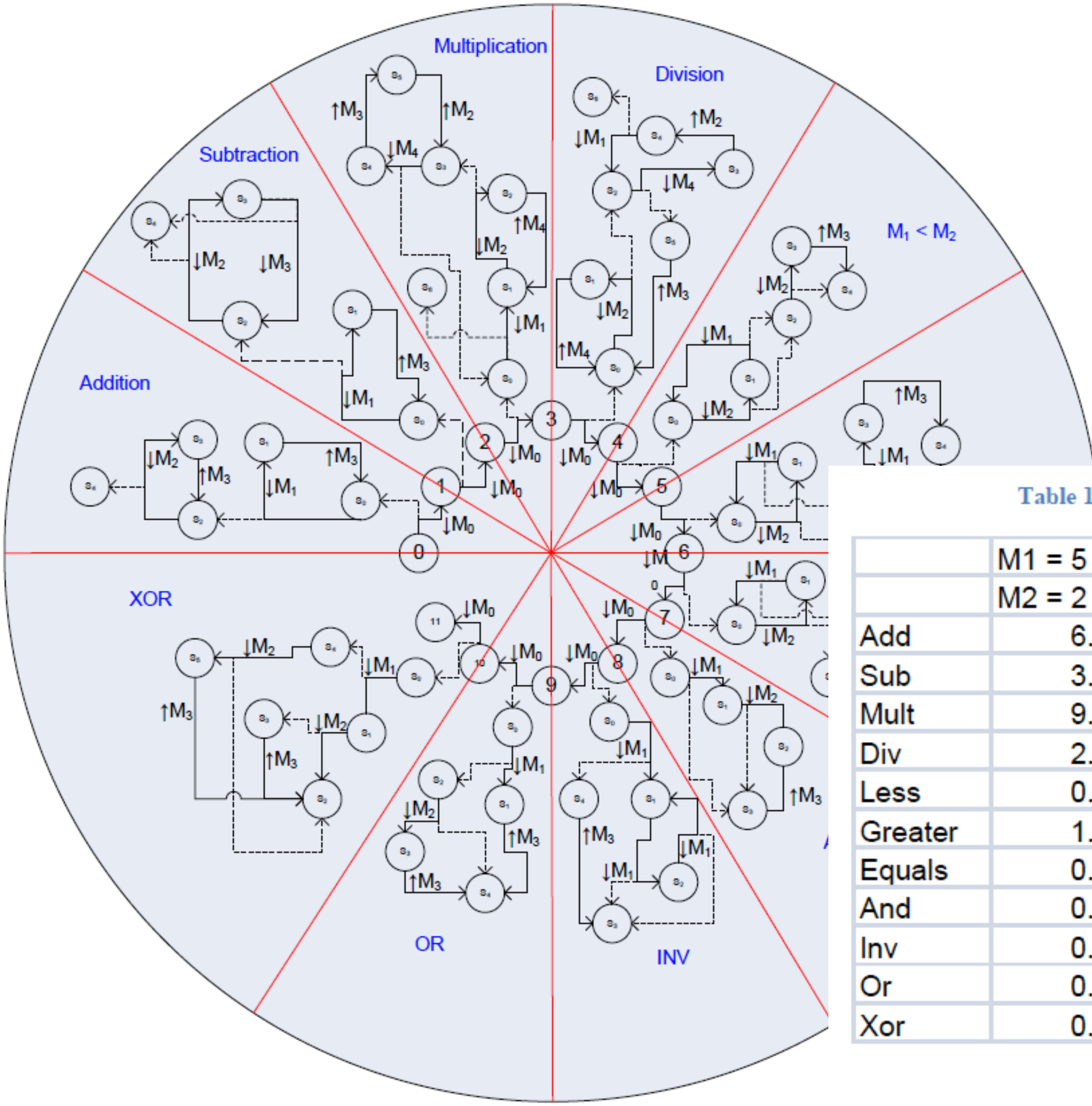


Figure 3: Entire BALU State Diagram

Table 1: Averaged Simulation Data

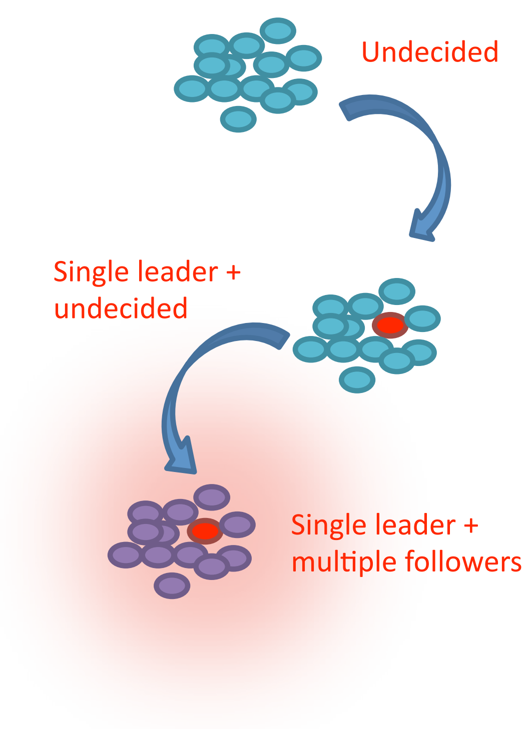
	M1 = 5	M1 = 2	M1 = 2	M1 = 10
	M2 = 2	M2 = 2	M2 = 5	M2 = 10
Add	6.83	3.88	6.99	19.39
Sub	3.04	0.17	0.22	1.09
Mult	9.02	3.56	8.76	87.9
Div	2.26	1.08	1.09	4.21
Less	0.14	0.09	1.28	1.54
Greater	1.08	0.14	0.24	2.24
Equals	0.01	0.87	0.26	1.02
And	0.96	1.02	1.25	1.89
Inv	0.13	0.17	0.15	0.33
Or	0.99	0.94	1.1	1.19
Xor	0.18	0.1	0.04	0.33

Zuyuan Zhang:
Term project, 2009.

Toward Synthetic Development

Leader Election

Electing a leader in a group of identical processes.

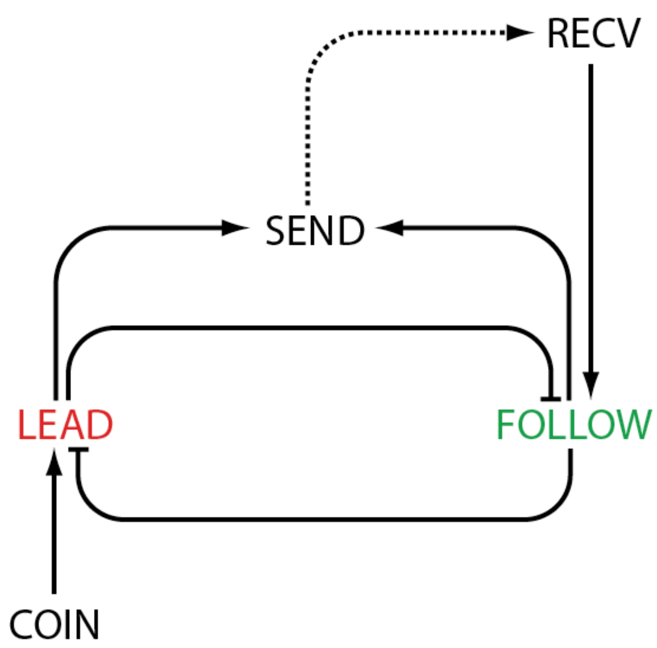


A simple approach:

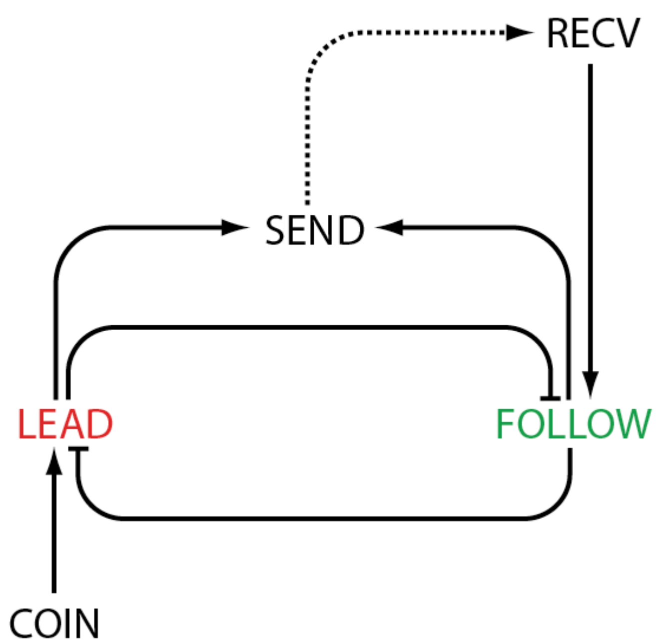
$$U \xrightarrow{k_1} L \qquad UL \xrightarrow{k_2} FL \qquad UF \xrightarrow{k_2} FF$$

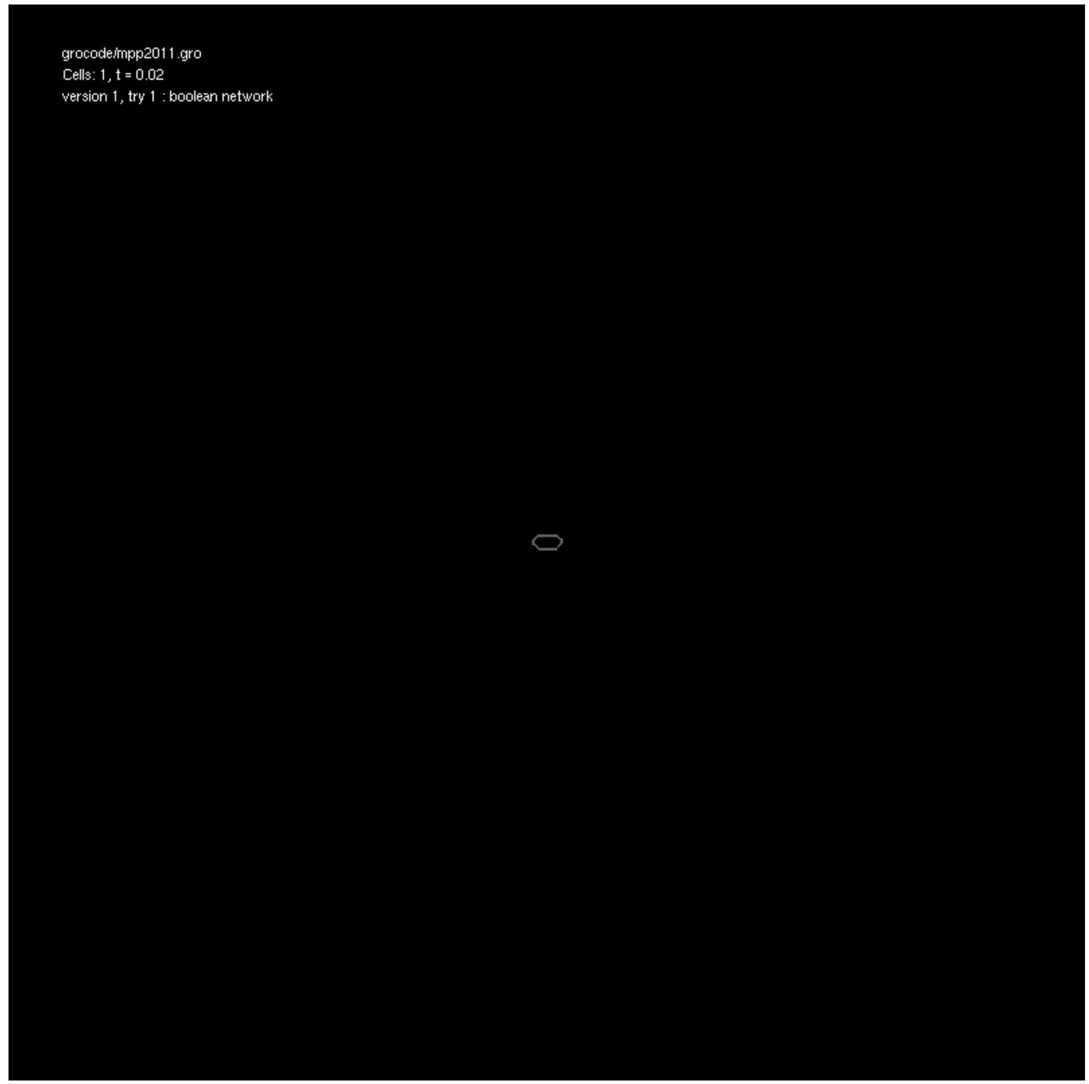
Better would be to include conflict resolution.

A Leader Election Circuit



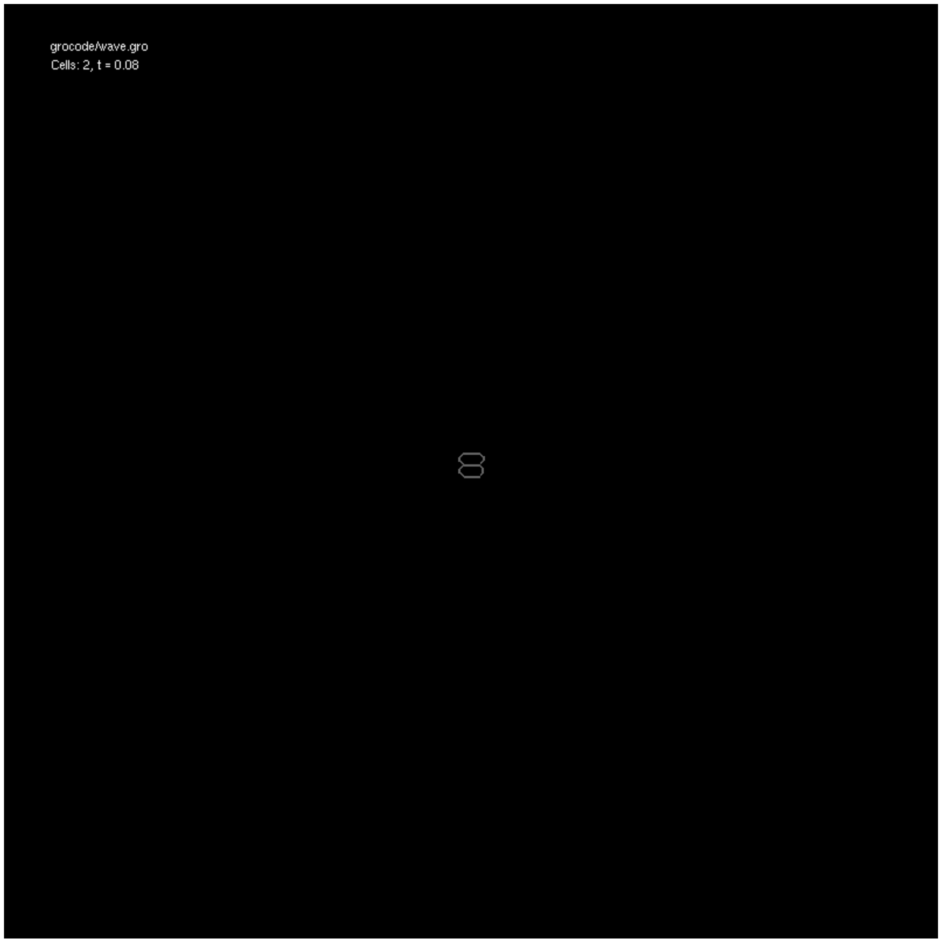
A Leader Election Circuit in gro





http://depts.washington.edu/soslab/gro

Another gro Program




```
ahl := signal ( 8, 6 );

program leader() := {

  gfp := 0;
  rfp := 0;
  p := [ t := 2.4 ]; // protected variable
  set ( "growth_rate", 0.00 );

  true : { p.t := p.t + dt }
  p.t > 2.5 : { emit_signal ( ahl, 100 ), p.t := 0 }

};

program follower() := {

  gfp := 0;
  rfp := 0;
  p := [ mode := 0, t := 0 ];
  set ( "growth_rate", 0.04 );

  p.mode = 0 & get_signal ( ahl ) > 0.01 :
    { emit_signal ( ahl, 100 ), p.mode := 1, p.t := 0 }
  p.mode = 1 : { p.t := p.t + dt }
  p.mode = 1 & p.t > 2.25 : { p.mode := 0 }

};

ecoli ( [ x:= 0, y:= 0, theta := 0 ], program leader() );
ecoli ( [ x:= 0, y:= 10, theta := 0 ], program follower() );
```

http://depts.washington.edu/soslab/gro

Outline

- Example Experimental Systems
- What Stochasticity Can Do
- Analytical Approaches
  - The Master Equation
  - Moment Dynamics
- Simulation Based Approaches
  - Simulation Methods
  - Approximate Abstraction/Refinement

Running Example:
Control of Gene Expression

Example:
Arithmetic

Software
Tools

Outline

- Example Experimental Systems
- What Stochasticity Can Do
- Analytical Approaches
    - The Master Equation
    - Moment Dynamics
- Simulation Based Approaches
    - Simulation Methods
    - Approximate Abstraction/Refinement

Running Example:
Control of Gene Expression

Example:
Arithmetic

Software
Tools

Questions

- **Convergence**: In probability, in mean and variance, via a Lyapunov Function.
- **Correctness**: Do individual trajectories do behave as expected?
- **Refinement**: What does it mean for a stochastic process to a refinement or coarse-graining of another stochastic process?

Probability vs. Time

Assumptions:

- The probability of when a given pair of molecules reacts in the next dt seconds is independent of time.
- A given molecule is equally likely to interact with every other molecule in the system.

Example: Consider a system with one A and one B and the reaction:

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

kdt is the probability that the reaction will occur in the next dt seconds. The two assumptions imply that the time of the reaction is distributed as an exponential random variable with p.d.f. and c.d.f.

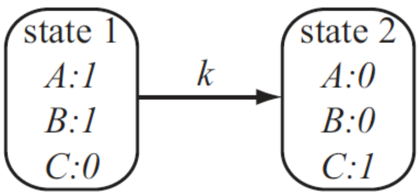
$$f(t) = ke^{-kt} \qquad F(t) = 1 - e^{-\alpha t}$$

Probability that the reaction has occurred by time t.

The Master Equation

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

Initially 1 A and 1 B:



$$p_1(t) = e^{-kt}$$
$$p_2(t) = 1 - e^{-kt}$$

Integrate

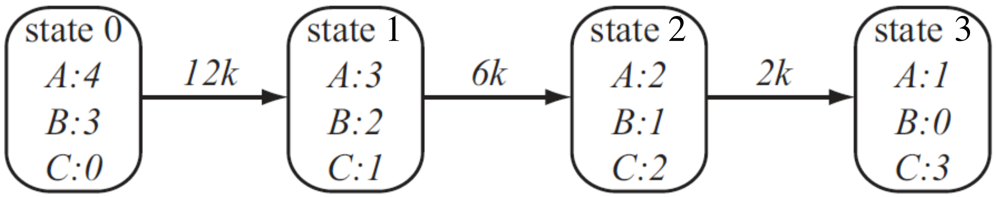
$$\dot{p}_1 = -kp_1$$
$$\dot{p}_2 = kp_1$$

Matrix Form

$$\dot{p} = Qp$$
$$Q = \begin{pmatrix} -k & 0 \\ k & 0 \end{pmatrix}$$

More Molecules

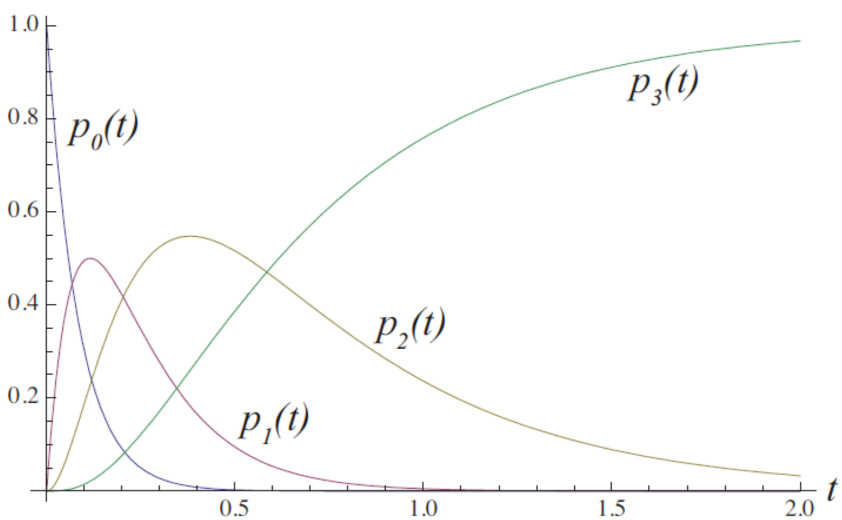
Initially 4 A's and 3 B's:

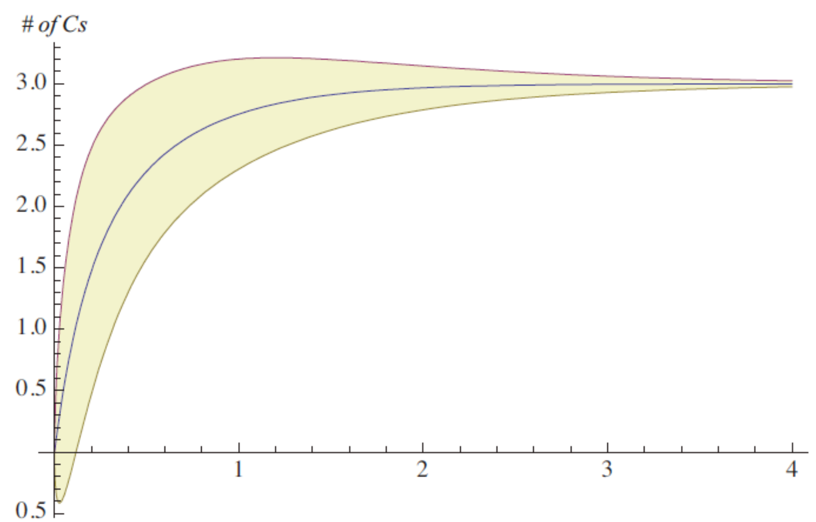


$$\begin{pmatrix} \dot{p}_0 \\ \dot{p}_1 \\ \dot{p}_2 \\ \dot{p}_3 \end{pmatrix} = \begin{pmatrix} -12k & 0 & 0 & 0 \\ 12k & 6k & 0 & 0 \\ 0 & -6k & -2k & 0 \\ 0 & 0 & 2k & 0 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

Note: easy to solve via

$$x(t) = e^{At}x(0)$$

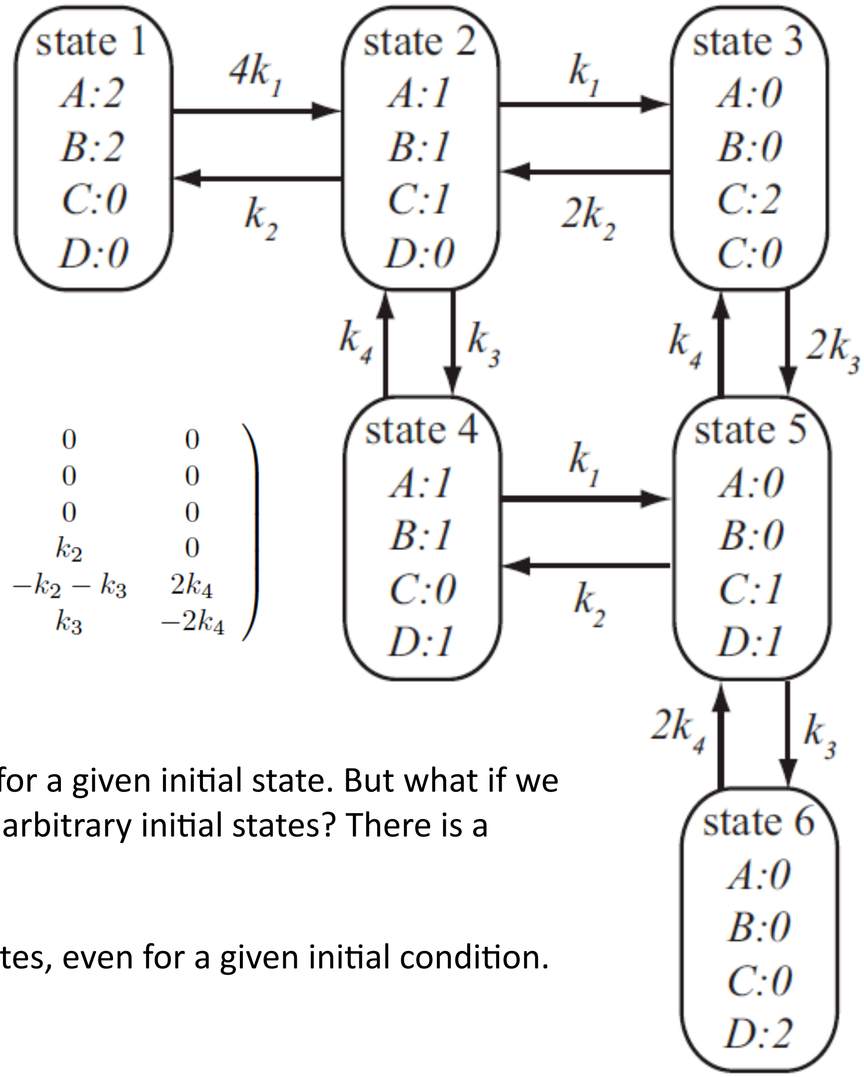




Another Example

$$A + B \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C$$

$$C \underset{k_4}{\overset{k_3}{\rightleftharpoons}} D$$

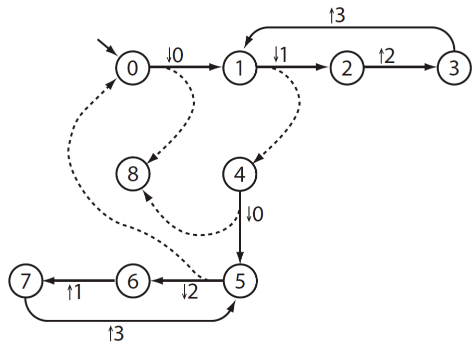


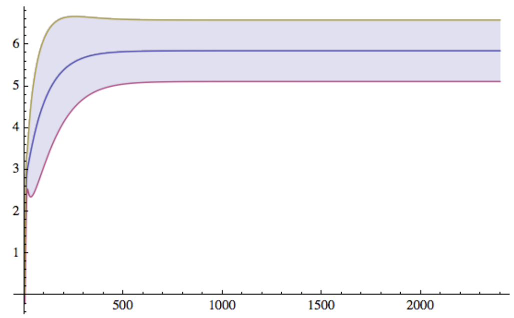
$$Q = \begin{pmatrix} -4k_1 & k_2 & 0 & 0 & 0 & 0 \\ 4k_1 & -k_1 - k_2 - k_3 & 2k_2 & k_4 & 0 & 0 \\ 0 & k_1 & -2k_2 - 2k_3 & 0 & 0 & 0 \\ 0 & k_3 & 0 & -k_1 - k_4 & k_2 & 0 \\ 0 & 0 & 2k_3 & k_1 & -k_2 - k_3 & 2k_4 \\ 0 & 0 & 0 & 0 & k_3 & -2k_4 \end{pmatrix}$$

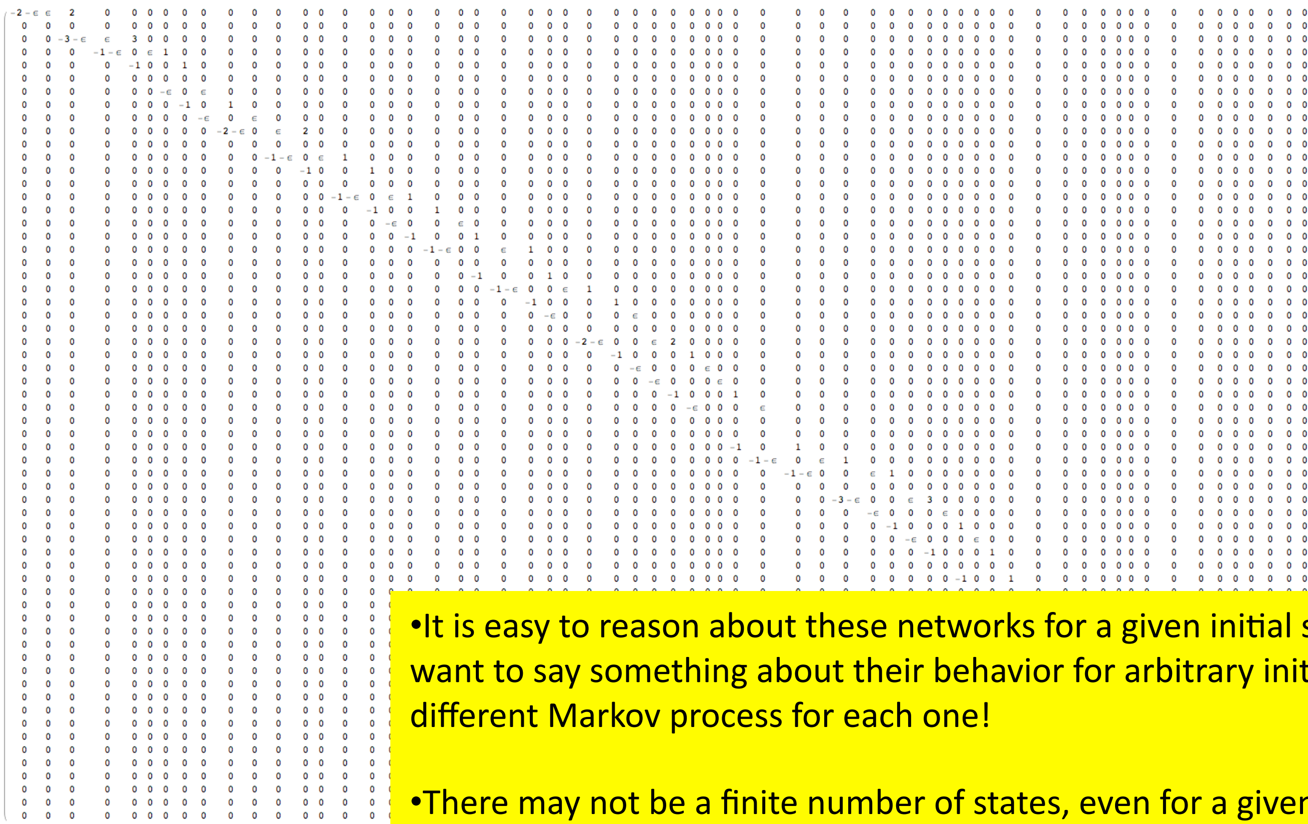
Note: It is easy to reason about this network for a given initial state. But what if we want to say something about its behavior for arbitrary initial states? There is a different Markov process for each one!

Also: There may not be a finite number of states, even for a given initial condition.

The Register Machine Example





$\dot{p} =$  p

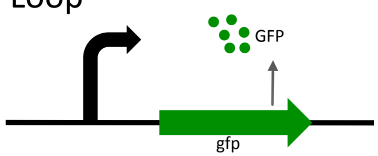
Reactions.m contains code to do this.

•It is easy to reason about these networks for a given initial state. But what if we want to say something about their behavior for arbitrary initial states? There is a different Markov process for each one!

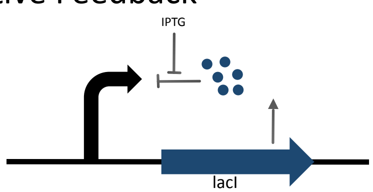
•There may not be a finite number of states, even for a given initial condition.

Running Example: Controlling Gene Expression

Open Loop



Negative Feedback



Other Control Schemes?

$$\emptyset \xrightarrow{r} X \xrightarrow{\beta} \emptyset$$

Goal: Control mean and variance of X.

r can be any function

- a constant
- a function of (the random variable) X
- a function of other species yet to be introduced
- etc.

Stationary Distribution of An Infinite System

$$\varnothing \xrightarrow{k_1} X \xrightarrow{k_2} \varnothing$$

$$\begin{array}{lllllll}
\dot{p}_0 & = & -k_1 p_0 & +k_2 p_1 & & & \\
\dot{p}_1 & = & k_1 p_0 & -(k_1+k_2)p_1 & +2k_2 p_2 & & \\
\dot{p}_2 & = & & k_1 p_1 & -(k_1+2k_2)p_2 & 3k_2 p_3 & \\
\dot{p}_3 & = & & & k_1 p_2 & -(k_1+3k_2)p_3 & 4k_2 p_3 \\
 & \vdots & & & & &
\end{array}$$

First Eqn:

$$p_1^* = \frac{k_1}{k_2} p_0^*$$

Second Eqn:

$$0 = -k_1 p_1^* + 2k_2 p_2^*$$

$$\Rightarrow \quad p_2^* = \frac{k_1}{2k_2} p_1^* = \frac{k_1^2}{2k_2^2} p_0^*$$

Sum of 1st n Eqns:

$$p_n^* = \frac{\alpha^n}{n!} p_0^* \quad \text{where } \alpha = k_1/k_2$$

Using the fact that p is a probability distribution:

$$\sum_{n=0}^{\infty} p_n^* = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} p_0^* = 1 \Leftrightarrow \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} = \frac{1}{p_0^*} \quad \Rightarrow \quad p_n^* = \frac{\alpha^n}{n!} e^{-\alpha}$$

$$\langle X \rangle^* = \sum n p_n^* = \alpha$$

$$\langle X^2 \rangle^* = \sum n^2 p_n^* = \alpha + \alpha^2$$

$$\mu_X = \frac{k_1}{k_2}$$

$$\sigma_X^2 = \frac{k_1}{k_2}$$

Note: Mean and variance can not be independently tuned by k_1. We need a better choice of control.

Solving Infinite Master Equations

- Although easy in simple cases, in general steady state distributions involve finding roots of high order polynomials symbolically.

- Some approaches:
  - Truncate the master equation (tends to work for numerical solutions)
  - Look at moments, instead of the full distribution
  - Simulate

Outline

- Example Experimental Systems
- What Stochasticity Can Do
- Analytical Approaches
    - The Master Equation
    - Moment Dynamics
- Simulation Based Approaches
    - Simulation Methods
    - Approximate Abstraction/Refinement

Example: Arithmetic

Software Tools

Running Example: Control of Gene Expression

Moment Dynamics

Say we have n reactions with rates $\lambda_i(x)$ and updates $x \mapsto \phi_i(x)$ for $i = 1$ to n.

Then

$$\frac{d}{dt}p(x) = \sum_r p(\phi_r^{-1}(x))\lambda_r(\phi_r^{-1}(x)) - p(x)\lambda_r(x)$$

is the Master Equation.

Let $\psi(x)$ be a test function with expected value

$$\langle\psi\rangle = \sum_x \psi(x)p(x).$$

Taking the derivative,

$$\begin{aligned}
\frac{d}{dt}\langle\psi\rangle &= \sum_x\sum_r \psi(x)p(\phi_r^{-1}(x))\lambda_r(\phi_r^{-1}(x)) - \sum_x\sum_r \psi(x)p(x)\lambda_r(x) \\
&= \sum_r\sum_y \psi(\phi_r(y))p(y)\lambda_r(y) - \sum_x\sum_r \psi(x)p(x)\lambda_r(x) \\
&= \sum_r\sum_x [\psi(\phi_r(x)) - \psi(x)]p(x)\lambda_r(x) \\
&= \langle\sum_r [\psi(\phi_r(X)) - \psi(X)]\lambda_r(X)\rangle \\
&\triangleq \langle L\psi\rangle
\end{aligned}$$

Example (No Control): $\varnothing \xrightharpoonup{u} X \xrightharpoonup{k} \varnothing$

$$
\begin{aligned}
\frac{d}{dt}\underset{\mu_1}{\langle X\rangle} &= \langle ((X+1)-X)u + ((X-1)-X)kX\rangle \\
&= \langle u\rangle - k\langle X\rangle \\
&= u - k\underset{\mu_1}{\langle X\rangle}
\end{aligned}
$$

$$
\begin{aligned}
\frac{d}{dt}\underset{\mu_2}{\langle X^2\rangle} &= \langle ((X+1)^2 - X^2)u + ((X-1)^2 - X^2)kX\rangle \\
&= \langle (2X+1u) + (-2X+1)kX\rangle \\
&= 2\langle Xu\rangle + \langle u\rangle - 2\langle X^2\rangle + k\langle X\rangle \\
&= u + (2u+k)\underset{\mu_1}{\langle X\rangle} - 2\underset{\mu_2}{\langle X^2\rangle}
\end{aligned}
$$

$$
\begin{pmatrix} \dot{\mu}_1 \\ \dot{\mu}_2 \end{pmatrix} = \begin{pmatrix} -k & 0 \\ 2u+k & -2k \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u
$$

$$
\mu_1 \rightarrow \frac{u}{k}
$$

$$
\sqrt{\mu_2 - \mu_1^2} = \sigma \rightarrow \sqrt{\frac{u}{k}}
$$

Note: Mean and variance can not be independently tuned by u. We need a better choice of control.

Example (Feedback): $\varnothing \xrightarrow{r-kX} X \xrightarrow{\beta} \varnothing$

r – kX is impossible to implement (a rate can't be negative).

But,

a) We are interested in the local behavior of the stationary distribution for smallish fluctuations and r – kX is the constant and linear part of whatever f(X) we do implement.

b) If u is non-linear, the moments don't close:

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \\ X_1^2 \\ X_2^2 \\ X_1 X_2 \\ X_1^3 \\ \vdots \end{pmatrix}$$

$$\dot{\mu} = A\mu + B$$

If rates are not unimolecular or constant, then some moments of order n will depend on higher order moments.

: This is also when the master difficult.

: Moment equations may still help.

: Various approximations exist (e.g. cumulant truncation).

Example: Feedback $\varnothing \xrightarrow{r-kX} X \xrightarrow{\beta} \varnothing$

Define test functions

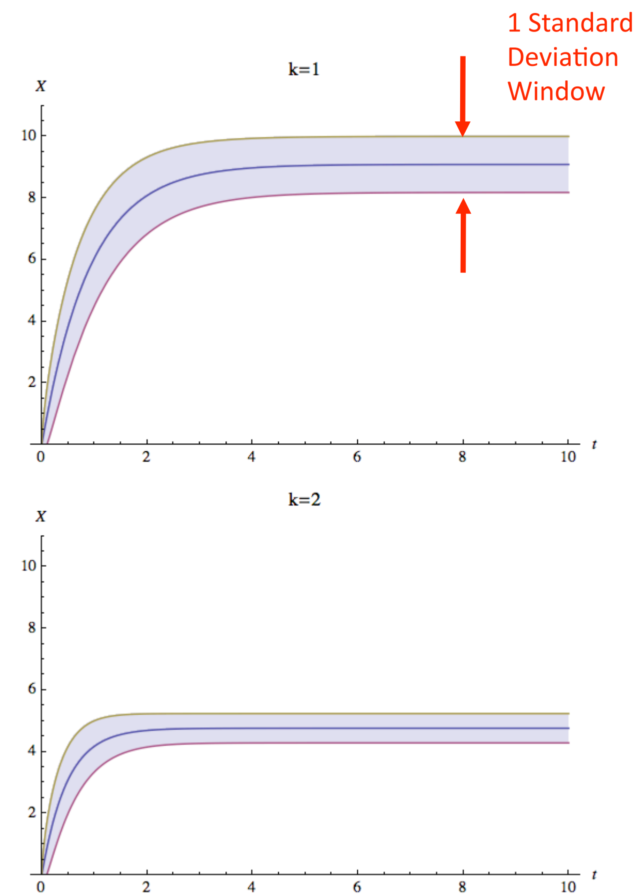
$$\mu_x = \langle X \rangle$$

$$\sigma_X^2 = \langle X^2 \rangle - \langle X \rangle^2$$

Use $\dfrac{d}{dt}\langle \psi \rangle = \langle L\psi \rangle$ to get

$$\frac{d}{dt}\mu_X = r - (k+\beta)\mu_X$$

$$\frac{d}{dt}\sigma_X^2 = r + (\beta - k)\mu_X - 2(\beta + k)\sigma_X^2$$



The stationary distribution satisfies:

$$\mu^* = \frac{r\beta}{(k+\beta)^2} \quad \text{and} \quad \kappa^* = \frac{r}{k+\beta}$$

Tunable, but
- mean sensitive to degradation rate
- variance coupled to r.

Aside: Mathematica Code

```
In[1]:= Import["/Users/ericklavins/Presentations/DNA17/Reactions.m"]

In[20]:= species = {"X"};
         reactions = {
            {"0", "X", r - k X},
            {"X", "0", β}
           };
         sys = {species, reactions};

In[33]:= cv = CumulantVector[sys]
         ce = CumulantDynamics[sys];
         ce // TableForm
         ss = Solve[SteadyState[ce], cv][[1]]
```

Out[33]= $\{\kappa_{XX}[t], \kappa_X[t]\}$

Out[35]//TableForm=

$$\kappa_X'[t] == r - k\,\kappa_X[t] - \beta\,\kappa_X[t]$$
$$\kappa_{XX}'[t] == r - k\,\kappa_X[t] + \beta\,\kappa_X[t] - 2\,k\,\kappa_{XX}[t] - 2\,\beta\,\kappa_{XX}[t]$$

Out[36]= $\left\{\kappa_{XX}[t] \to \frac{r\,\beta}{(k+\beta)^2},\ \kappa_X[t] \to \frac{r}{k+\beta}\right\}$

```
In[52]:= sol1 = NDSolve[(ce /. {r → 10, k → 1, β → 0.1}) ⋃ {κ"XX"[0] == 0, κ"X"[0] == 0}, cv, {t, 0, 10}];
         g1 = MeanVarPlot["X", sol1, 0, 10];
```

Reactions.m includes:

- Mass action kinetics
- Markov Processes and Master Equations
- Gillespie Simulations
- Moment and Cumulant Dynamics Analysis

Back to $\varnothing \xrightarrow{u} X \xrightarrow{k} \varnothing$

Idea: Proportional-Integral Control

$$u = \gamma Z - kX$$

$$\dot{Z} = r - X$$

This is now a continuous integrator.
At steady state, z=X.

- Z_{off} in large supply
- Reverse rate saturates

$$Z_{off} \underset{X}{\overset{\tilde{r}}{\rightleftharpoons}} Z$$

$$r \longrightarrow Z \rightleftharpoons X$$

$$h(X,Z) = \frac{vZ^m}{(K_Z + Z^m)(K_X + X^n)} \approx h(\mu_X^*, \mu_Z^*) + \left.\frac{\partial h X, Z}{\partial X}\right|_{\mu=\mu^*} (X - \mu_X^*) + \left.\frac{\partial h X, Z}{\partial Z}\right|_{\mu=\mu^*} (Z - \mu_Z^*) = \alpha + \gamma Z - kX.$$

What about the mixed discrete/
continuous system?

Mixed Continuous / Discrete

Suppose we have a concurrent continuous process

$$\dot{z} = f(X_1, ..., X_n, z).$$

Then the L can be extended to

$$L\psi = \frac{\partial \psi}{\partial z} f(X_1, ..., X_n, z) + \sum_i (\psi_{new} - \psi) k_i$$

$$\frac{d}{dt}\langle \psi \rangle = \langle L\psi \rangle$$ still works!

Proportional-Integral Control

$$\varnothing \xrightharpoonup{u} X \xrightharpoonup{k} \varnothing$$

$$\dot{z} = X - r$$

$$u = h[-k_P(X-r) - k_I z]$$

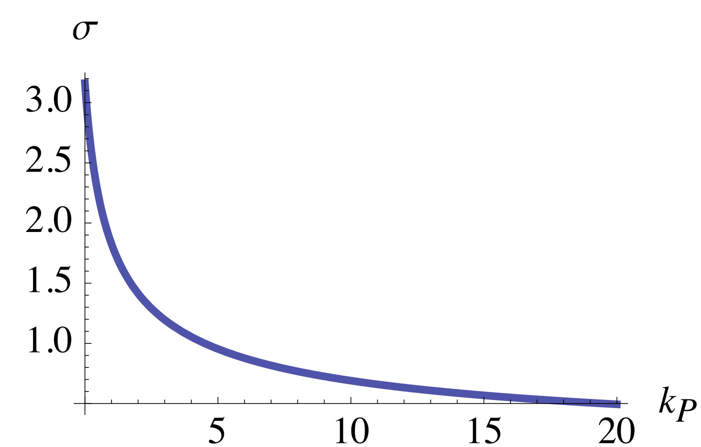
$$\frac{d}{dt}\begin{pmatrix} \langle X \rangle \\ \langle Z \rangle \\ \langle X^2 \rangle \\ \langle XZ \rangle \\ \langle Z^2 \rangle \end{pmatrix} = \begin{pmatrix} -k-k_P & -k_I & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ k-k_P+2k_Pr & -k_I & -2k-2k_P & -2k_I & 0 \\ -r & k_Pr & 1 & -k-k_P & -k_I \\ 0 & -2r & 0 & 2 & 0 \end{pmatrix}\begin{pmatrix} \langle X \rangle \\ \langle Z \rangle \\ \langle X^2 \rangle \\ \langle XZ \rangle \\ \langle Z^2 \rangle \end{pmatrix} + \begin{pmatrix} k_P \\ -1 \\ k_P \\ 0 \\ 0 \end{pmatrix} r$$

- Stable (eigenvalues in left half plane) = convergence in mean and variance.
- The mean value of X converges to r (insensitively).

$$\langle \dot{Z} \rangle = \langle X \rangle - r = 0 \Rightarrow \langle X \rangle^* = r$$

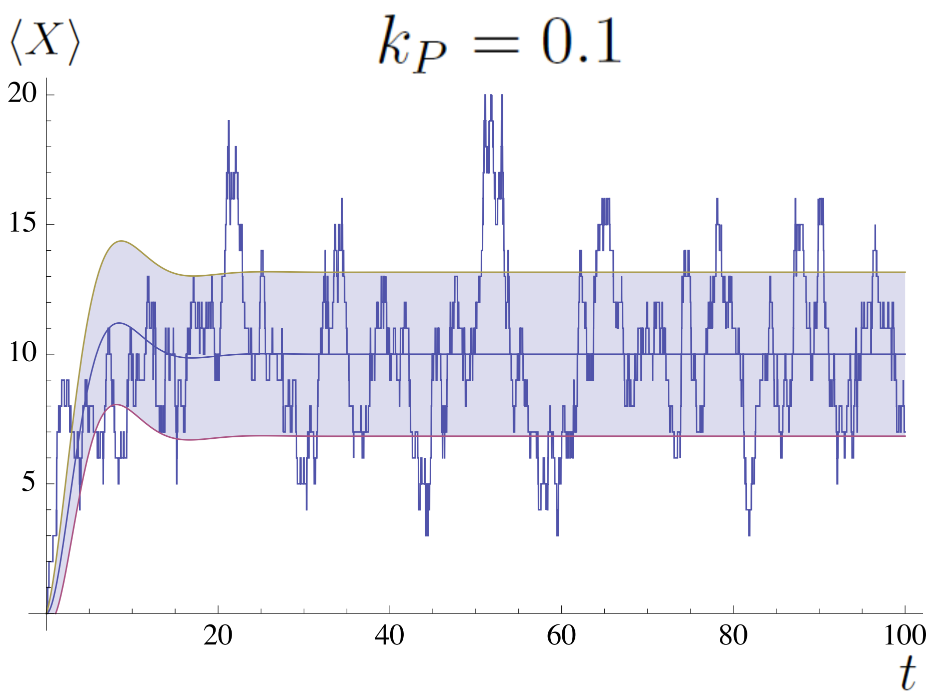
- The steady state standard deviation is tunable via k_P

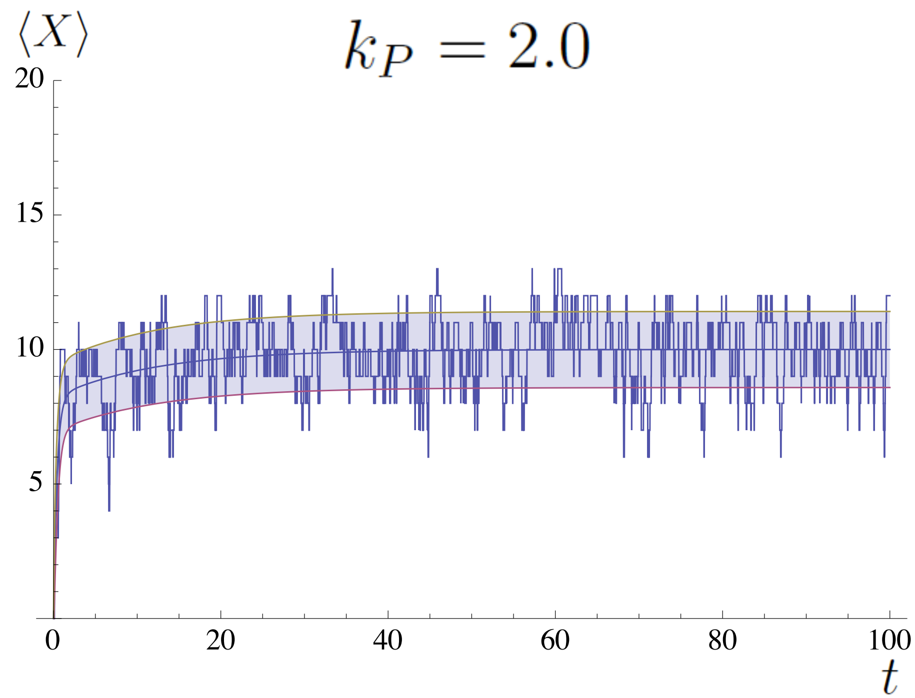
$$\sqrt{\langle X^2 \rangle - \langle X \rangle^2} = \sigma \to \sqrt{\frac{kr + kr^2 + k_Pr^2}{k+k_P} - r^2}$$



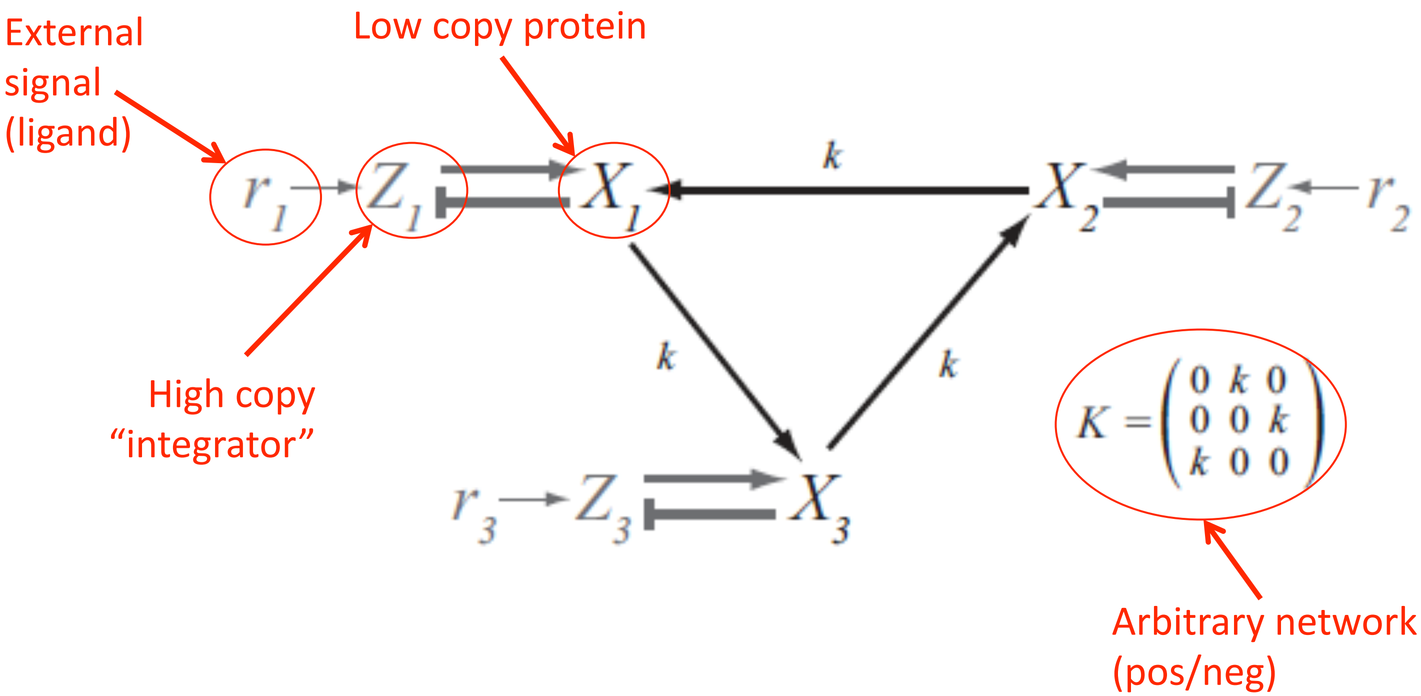
Simulations

$$\varnothing \xrightarrow{u} X \xrightarrow{k} \varnothing$$
$$\dot{z} = X - r$$
$$u = h[-k_P(X - r) - k_I z]$$





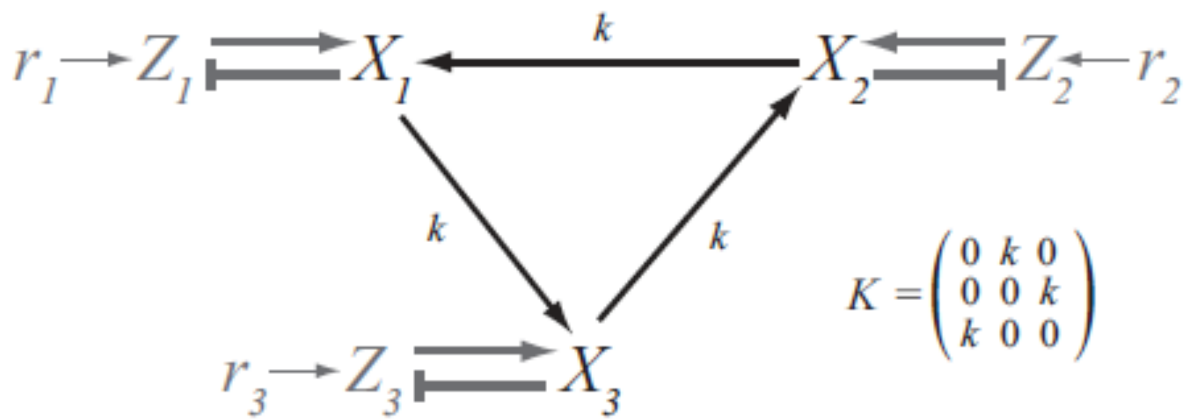
A Class of Network Structures



PI Control in the literature

- Alon's chapter on chemotaxis
- Napp, Burden and Klavins: Control of Stochastic Robotics
- Two component systems in general

A Class of Network Structures



Gene expression: (Discrete) $\varnothing \xrightarrow{u_i(X,Z)} X_i \xrightarrow{\beta_i} \varnothing$

Regulation: $u_i(X, Z) = \gamma_i Z_i - \sum_{j=1}^{n} k_{ij} X_j$

Integration: (Continuous) $\dot{Z}_i = r_i - X_i$

A Stochastic Hybrid System

… with closed moment dynamics.

Moments

Group the means and moments into vectors and matrices

$$\mu = \begin{pmatrix} \mu_X \\ \mu_Z \end{pmatrix} \triangleq \begin{pmatrix} \langle X \rangle \\ \langle Z \rangle \end{pmatrix} \quad \text{and} \quad M \triangleq \begin{pmatrix} \langle X X^T \rangle & \langle X Z^T \rangle \\ \langle Z X^T \rangle & \langle Z Z^T \rangle \end{pmatrix}$$

$$\kappa = \begin{pmatrix} \kappa_{XX^T} & \kappa_{XZ^T} \\ \kappa_{ZX^T} & \kappa_{ZZ^T} \end{pmatrix} \triangleq M - \mu\mu^T$$

Group the parameters

P	$\triangleq$	$\mathrm{diag}(\beta_1, ..., \beta_n)$	Degradation
Γ	$\triangleq$	$\mathrm{diag}(\gamma_1, ..., \gamma_n)$	Integrator gain (tunable)
K	$\triangleq$	$\{k_{ij}\}$	Network (tunable)
r	$\triangleq$	$(\ r_1\ ...\ r_n\ )^T.$	Reference inputs

Moment Dynamics

Use the extended generator to get mean dynamics

$$\frac{d}{dt}\begin{pmatrix}\mu_X \\ \mu_Z\end{pmatrix} = \begin{pmatrix}-P-K & \Gamma \\ -I & 0\end{pmatrix}\begin{pmatrix}\mu_X \\ \mu_Z\end{pmatrix} + \begin{pmatrix}0 \\ I\end{pmatrix} r$$

$$\dot{\mu} = A\mu + Br$$

And the second moment dynamics

$$\dot{M} = AM + MA^T + C(\mu)$$

$$C(\mu) = \begin{pmatrix}\mathrm{diag}(\Gamma\mu_Z + (P-K)\mu_X) & \mu_X r^T \\ r\mu_X^T & \mu_Z r^T + r\mu_Z^T\end{pmatrix}$$

Properties

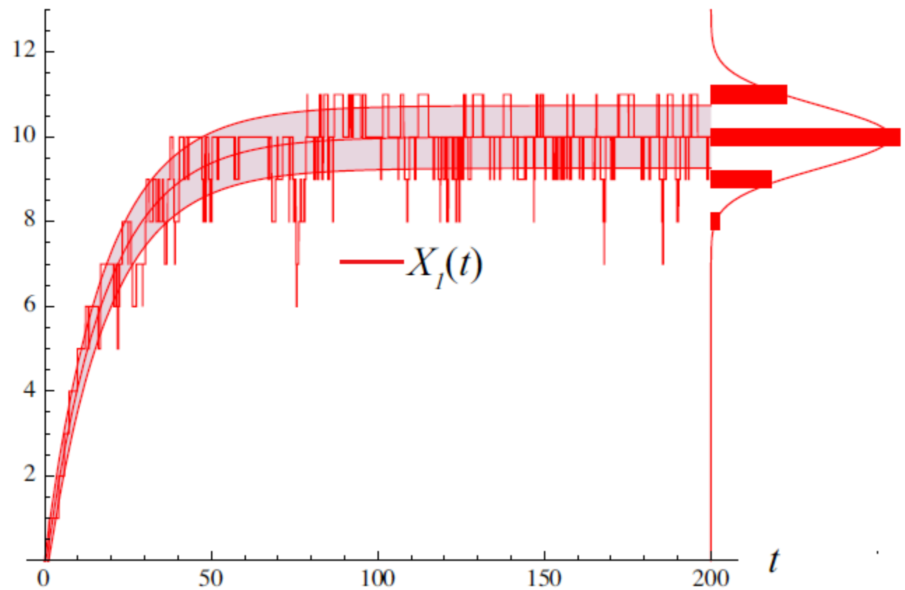
$$\dot{\mu} = A\mu + Br$$
$$\dot{M} = AM + MA^T + C(\mu)$$

Theorem 1: The network converges in mean and variance if and only if A is Hurwitz.

Theorem 2: The unique steady state mean μ_X^* is r and is insensitive to K, Γ, and P.

Theorem 3: The steady state covariance matrix in X can be placed arbitrarily. That is, if W is positive definite, then K and Γ can be found so that in steady state $\kappa_{XX^T}^* = W$.

Example: One Gene



c) Proportional/integral control

$$\varnothing \xrightarrow{\gamma Z - kX} X \xrightarrow{\beta} \varnothing$$

$$\dot{Z} = r - X$$

$$\mu^* = r \quad \text{and} \quad \kappa^* = \frac{r\beta}{k+\beta}$$

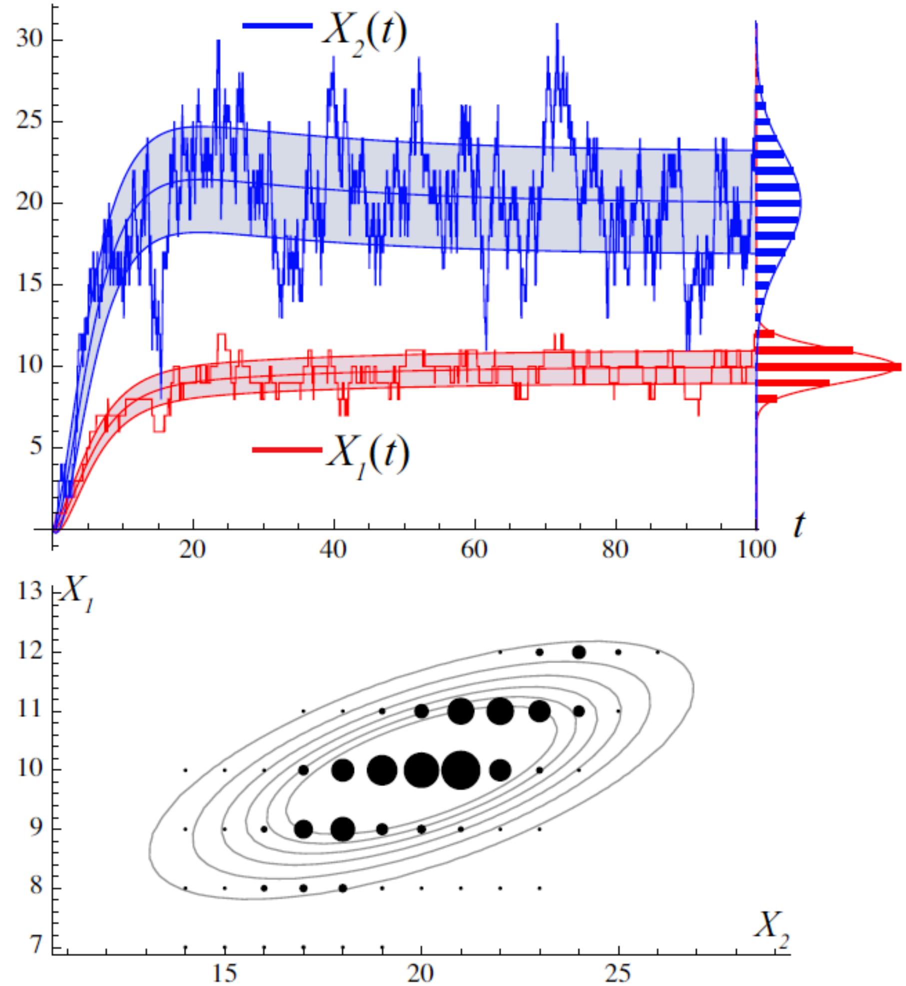
Mean and variance independently tunable

Example: Two Genes

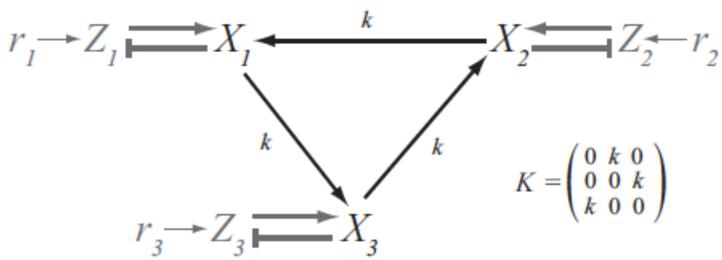
Specification

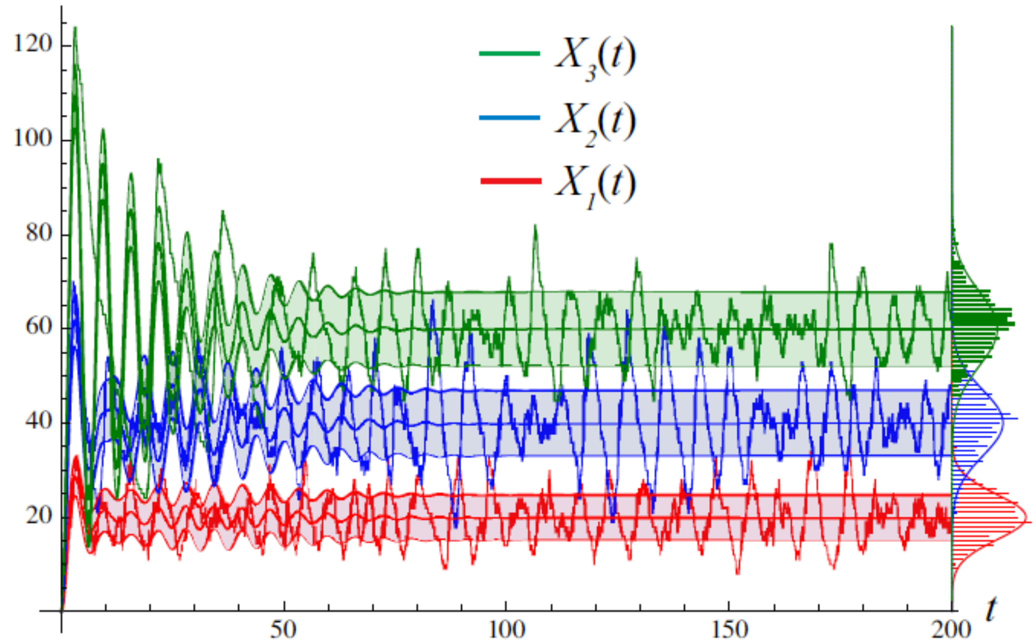
$$\mu^* = r = \begin{pmatrix} 10 \\ 20 \end{pmatrix}$$

$$\kappa_{XX^T} = \begin{pmatrix} 1 & 2 \\ 2 & 10 \end{pmatrix}$$



Example: Excitable Oscillator





Ensemble dynamics: Damped oscillator ($-a_i+b_i j$, $-a_i-b_i j$).

Stochastic dynamics: Sloppy oscillations with specific means.

Example #2

$$
\begin{tikzcd}
\varnothing \arrow[r, "k_1"] & X \arrow[r, shift left, "u"] \arrow[d, "k_4"'] & Y \arrow[l, shift left, "k_2"] \arrow[r, "k_3"] & \varnothing \\
& \varnothing & &
\end{tikzcd}
$$

<u>Full state feedback controller with an integrator</u>

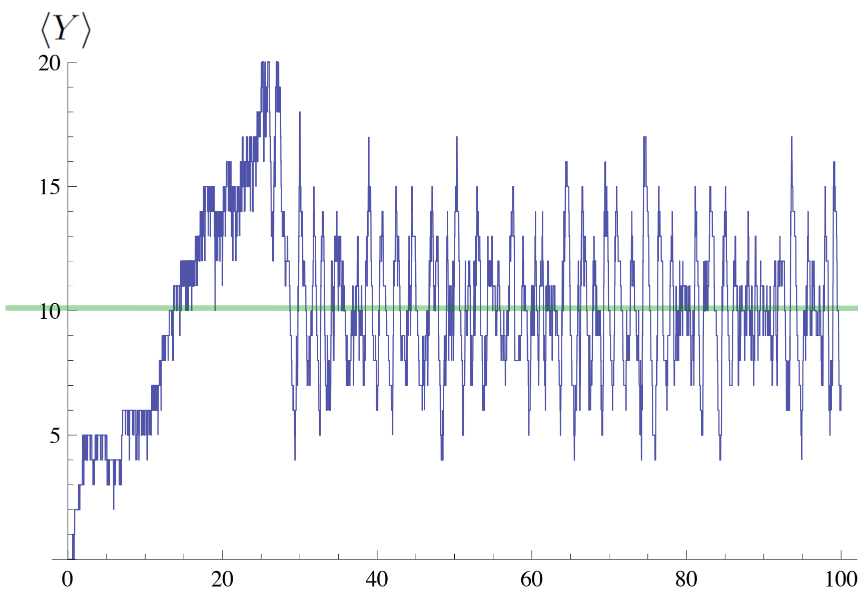
$$\dot{z} = Y - r$$

$$u = h[-k_{Px}X - k_{Py}Y - z]$$

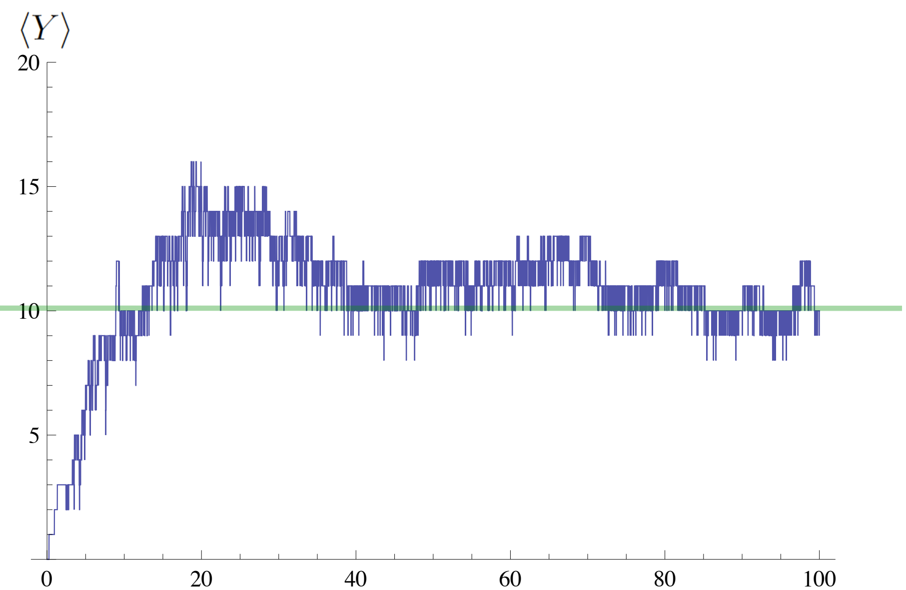
$$k_{Px} < 0, \quad k_{Py} > 0$$

Example #2

$$\begin{tikzcd} \varnothing \arrow[r, "k_1"] & X \arrow[r, shift left, "u"] \arrow[d, "k_4"'] & Y \arrow[l, shift left, "k_2"] \arrow[r, "k_3"] & \varnothing \\ & \varnothing & & \end{tikzcd}$$



k_{Px}=-0.01, k_{Py}=0.01



k_{Px}=-1, k_{Py}=1

Open Moments

$$
\begin{tikzcd}
\varnothing \arrow[r, "k_1"] & X \arrow[r, shift left, "u"] \arrow[d, "k_4"'] & Y \arrow[l, shift left, "k_2"] \arrow[r, "k_3"] & \varnothing \\
 & \varnothing & &
\end{tikzcd}
$$

<u>Steady State Moment Equations Give</u>

$$\langle \dot{Z} \rangle = \langle Y \rangle - r = 0 \;\not\Rightarrow\; \langle Y \rangle^* = r$$

<u>Example Second Moment</u>

$$\frac{d}{dt}\langle YZ \rangle = -r\langle Y \rangle + \langle Y^2 \rangle - (k_2 + k_3)\langle Z \rangle - k_{Px}\langle X^2 Z \rangle - k_{Py}\langle XYZ \rangle - k_I\langle XZ^2 \rangle$$

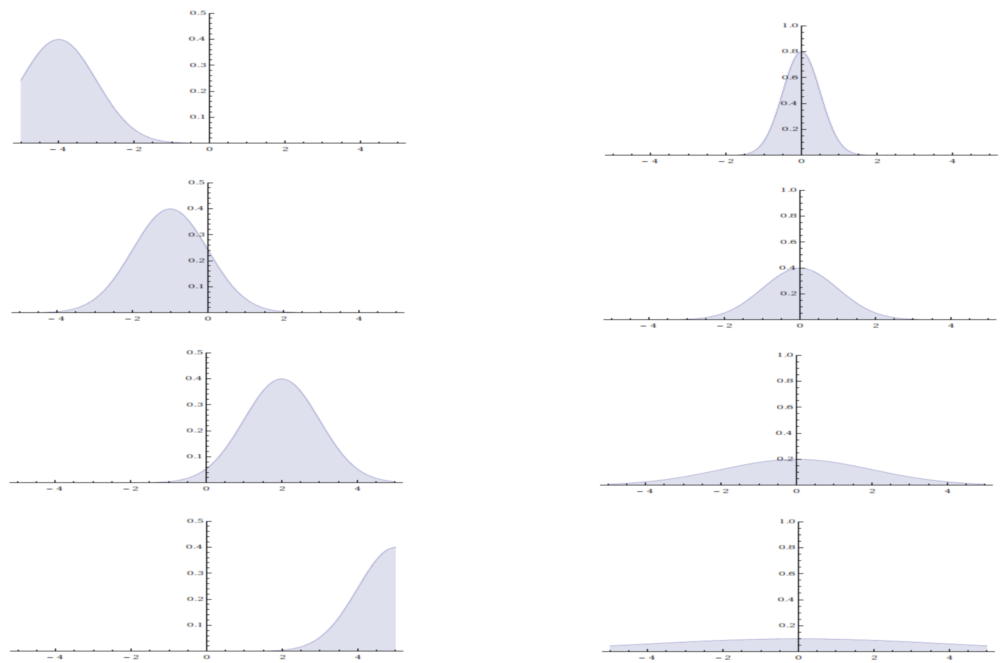
Second Order

Third Order ☹

Idea: Approximate Higher Order Moments? => No general results.

Proving Convergence

$$\langle \dot{Z} \rangle = \langle X \rangle - r = 0 \Rightarrow \langle X \rangle^* = r$$ *WHEN THE SYSTEM IS ERGODIC*



If the moments are closed, you can check for a stable steady state or just reason about the mean and variance.

If not, some other argument must be used.

Lyapunov Criterion for Markov Processes

Theorem (Meyn): If for some compact region C and positive constant ε, there exists a positive radially unbounded function $V(q,x)$ such that

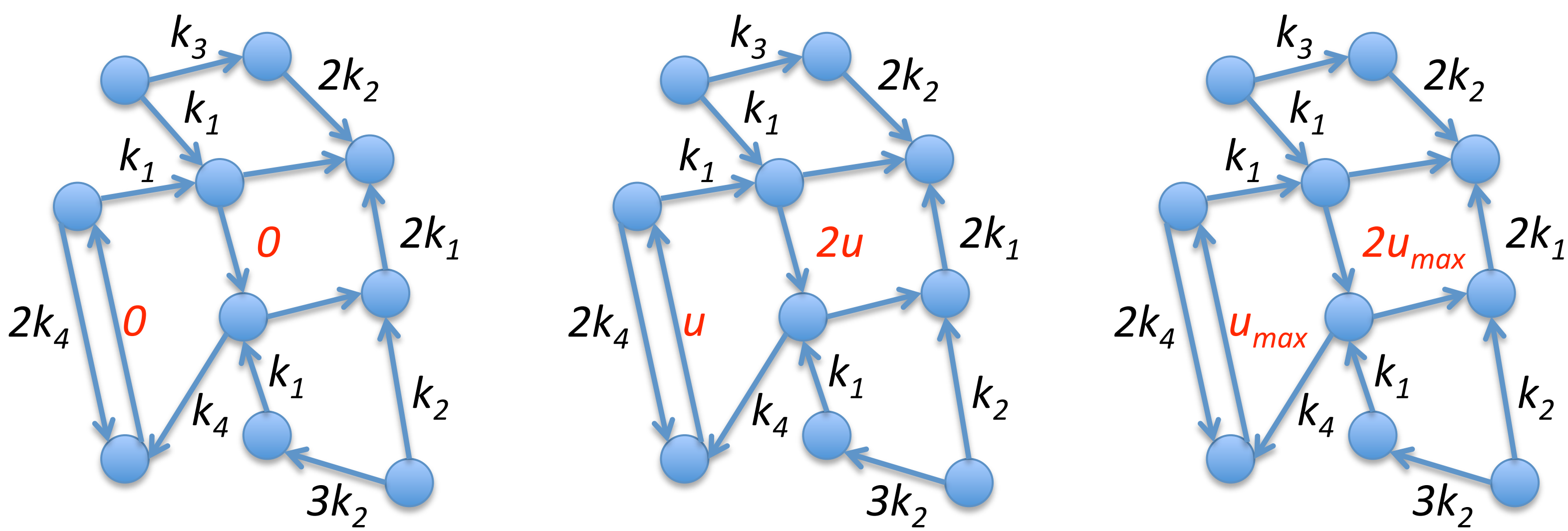
$$\mathcal{L}V(q,x) \leq -\epsilon \qquad \forall (q,x) \notin C$$

then the process is ergodic.

I.e., the expected value of V decreases outside of C.

S. Meyn, R. L. Tweedie. *Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes*. Advances in Applied Probability, 25(3):518-48, 1993. Thm 5.1.

The Controllable Region



A_{min}
p_{min}
y_{min}

A_{max}
p_{max}
y_{max}

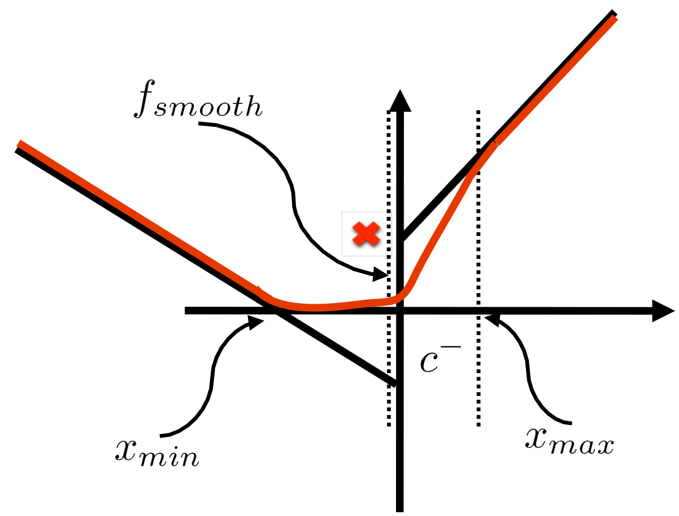
Can only reasonably expect to achieve r in $[y_{min}, y_{min}]$.

Integral Control Works for <u>any</u> SCRN

Theorem (Napp and Klavins): Suppose steady state distributions p_{min} and p_{max} for constant inputs u_{min} and u_{max} respectively and suppose that

1. $\dot{z} = Y - r$;
2. $u = h[-k_i z]$ with $k_i > 0$
3. $p_{min} \leq r \leq p_{max}$.

Then $\langle Y \rangle \to r$ with finite variance.



$$\widetilde{V}(q,x) = \begin{cases} -x + c^-(q), & x \leq x_{min} \\ f_{smooth}(q,x), & x_{min} < x \leq x_{max} \\ x + c^+(q), & x < x_{max} \end{cases}$$

Napp&Klavins, 2010

Outline

- Example Experimental Systems
- What Stochasticity Can Do
- Analytical Approaches
    - The Master Equation
    - Moment Dynamics
- **Simulation Based Approaches**
    - Simulation Methods
    - Approximate Abstraction/Refinement

Running Example:
Control of Gene Expression

Example:
Arithmetic

Software
Tools

Outline

- Example Experimental Systems
- What Stochasticity Can Do
- Analytical Approaches
  - The Master Equation
  - Moment Dynamics
- Simulation Based Approaches
  - **Simulation Methods**
  - Approximate Abstraction/Refinement

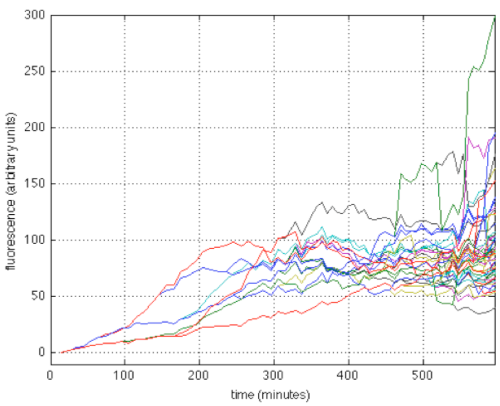
Running Example:
Control of Gene Expression

Example:
Arithmetic

Software
Tools

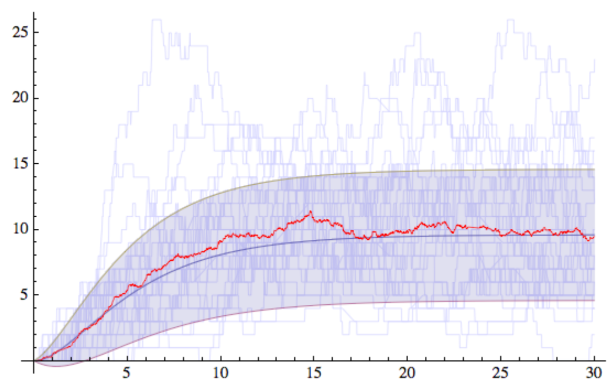
Data from Simulations and Experiments

What really happens in the cell



$$\varnothing \xrightarrow{r} X \xrightarrow{\beta} \varnothing$$

A simple model



What about:

- A complex model?
- A refinement?
- An implementation?
- …

Plan: Reason about systems based on the data they produce.

Simulation Approaches

The Stochastic Simulation Algorithm (Gillespie's SSA)

Next reaction

$$P[t,j] = k_{i,j} e^{-K_i t}$$

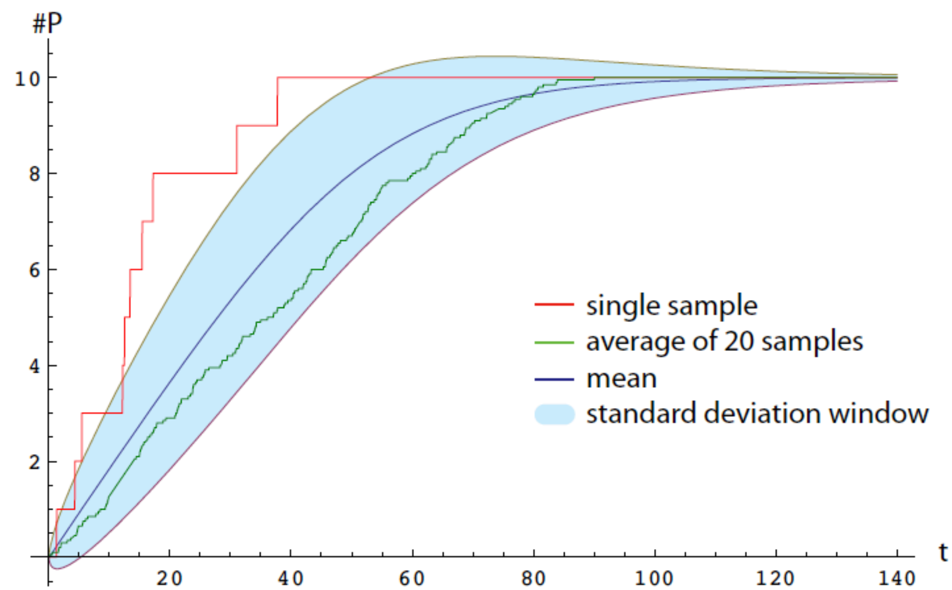
$$p_{i,j} = \int_0^\infty k_{i,j} e^{-K_i t} dt = \frac{k_{i,j}}{K_i}$$

Time of the next reaction

$$\int_0^t K_i e^{-K_i t \tau} d\tau = 1 - e^{-K_i t} = r \in [0,1]$$

$$\sum_{j=1}^{N} k_{i,j} e^{-K_i t} = K_i e^{-K_i t}$$

$$t = \frac{1}{K_i} \ln \frac{1}{1-r}$$



1. Choose an initial condition v equal to some vector of the copy numbers of the species in the reaction network.
2. Set $t = 0$.
3. For each reaction ρ applicable in v, determine the rate k_v.
4. Choose the next reaction via Equation 8.6.
5. Choose the Δt via Equation 8.7 and set t to $t + \Delta t$.
6. Goto 3.

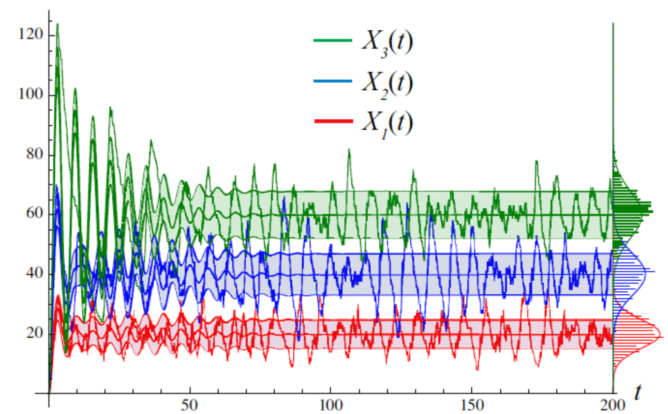
Simulation Approaches

Plain Old Euler Integration

- •With mixed discrete / continuous systems, the SSA doesn't directly work.
- •And there is diminishing return for systems with many reactions.

Choose a timestep δ such that $\delta\lambda_{max} < 1$ for the largest rate λ_{max} in your system.

- Enabled reactions $1, ..., N(t)$ with rates $\lambda_1, ..., \lambda_{N(t)}$.
- $\mu_i = \sum_{j<i} \lambda_i$.
- Choose $r \in [0, 1]$.
- Fire the reaction i such that μ_i is the larg exists. Otherwise, do nothing.
- $x(t + \delta) = x(t) + \delta f(x, q)$.
- $t \to t + \delta$.



gro Simulations

P = {
  $guard_1$:$command_1$
  ...
  $guard_n$:$command_n$
}

•Growth is continuous.

•Signaling is continuous (finite element sim).

•Physics via Chipmunk (which takes dt as an argument at each step).

•Guards may have rand(0,1)<0.25 evaluated at each iteration.

Outline

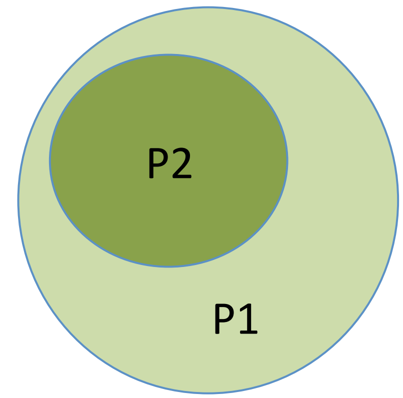
- Example Experimental Systems
- What Stochasticity Can Do
- Analytical Approaches
    - The Master Equation
    - Moment Dynamics
- Simulation Based Approaches
    - Simulation Methods
    - Approximate Abstraction/Refinement

Example: Arithmetic

Software Tools

Running Example: Control of Gene Expression

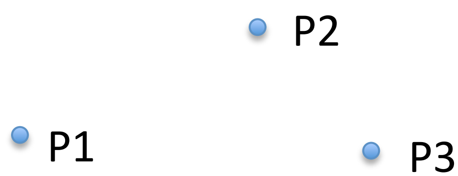
Approximation



In computer science, non-determinism has nice definitions for abstraction, refinement, implementation, simulation, etc.

For stochastic processes, what does it mean for one process to be an abstraction of another? A refinement? A coarse graining?

Approximate Bisimulation: Turns bismulation into a metric on processes. Distance zero means bismilar. Destance epsilon means close.



W(P1,P3) > W(P2,P3)

Comparing Stochastic Behaviors

Let f be a metric on the spaceof trajectories Ω.

For any two probability distributions P_1 and $P2$ on Ω, the *Wasserstein Metric* is defined by

$$W(P_1, P_2) = \inf_{Q \in J(P_1, P_2)} \int_{\Omega \times \Omega} f(\omega, \eta) dQ(\omega, \eta).$$

Hard to compute!

When Ω is finite, finding W amounts to solving the linear program

$$\text{Minimize} \quad \sum_i \sum_j f(\omega_i, \eta_j) Q_{i,j}$$

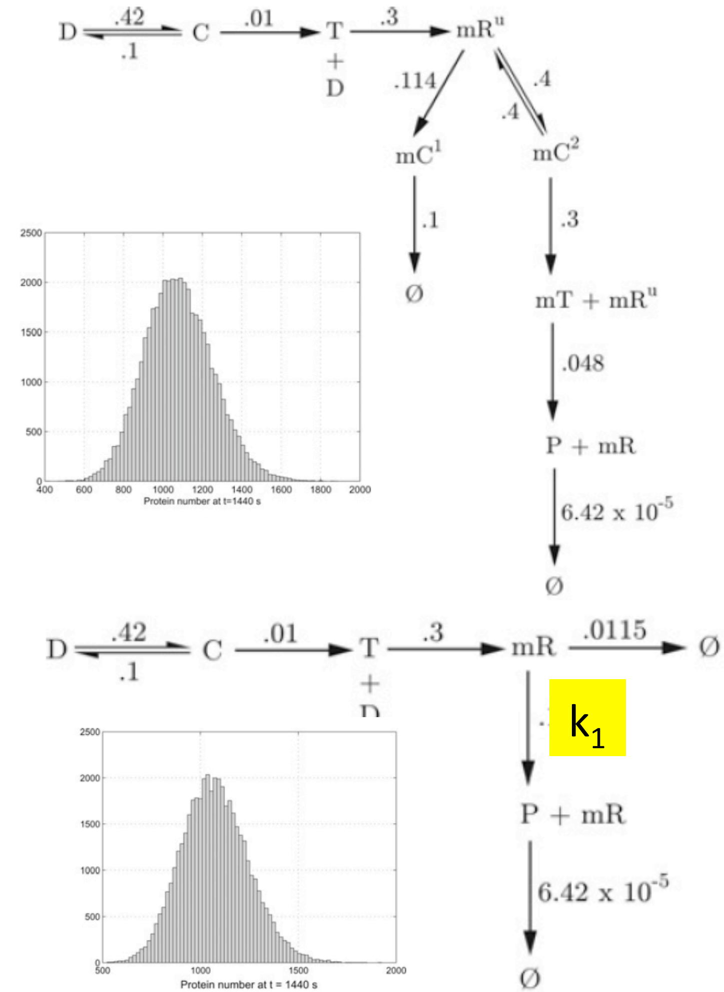
$$\text{Subject to} \quad \sum_j Q_{i,j} = \frac{1}{n}$$

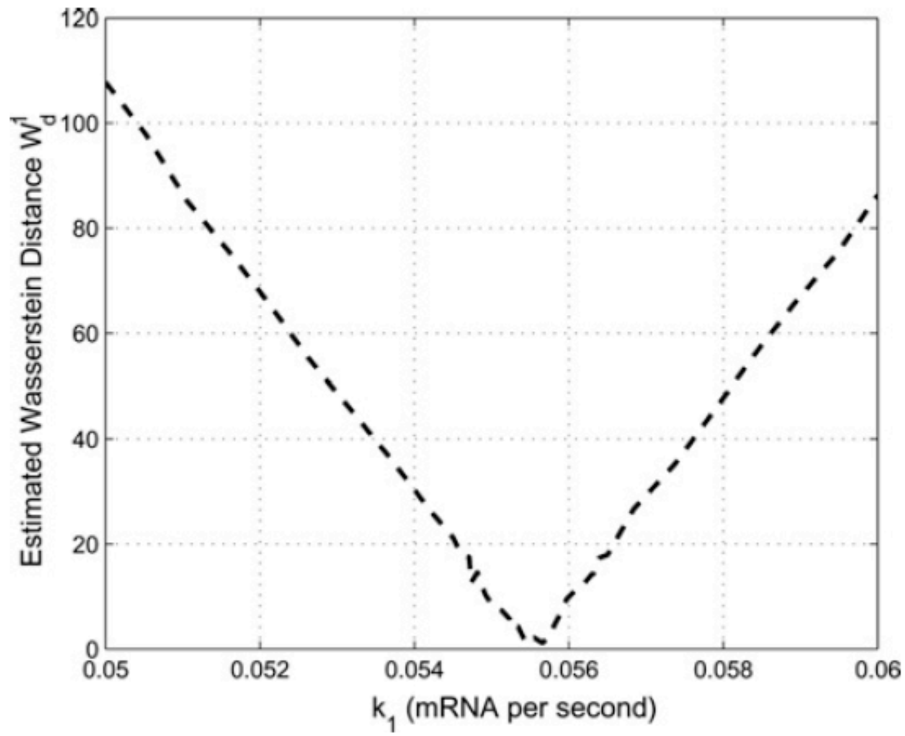
$$\sum_i Q_{i,j} = \frac{1}{n}$$

$$Q_{i,j} \geq 0$$

Thorsley and Klavins, "Approximating stochastic biochemical processes with Wasserstein pseudometrics", *IET Systems Biology*, June 2009.

Example: Abstracting Gene Expression



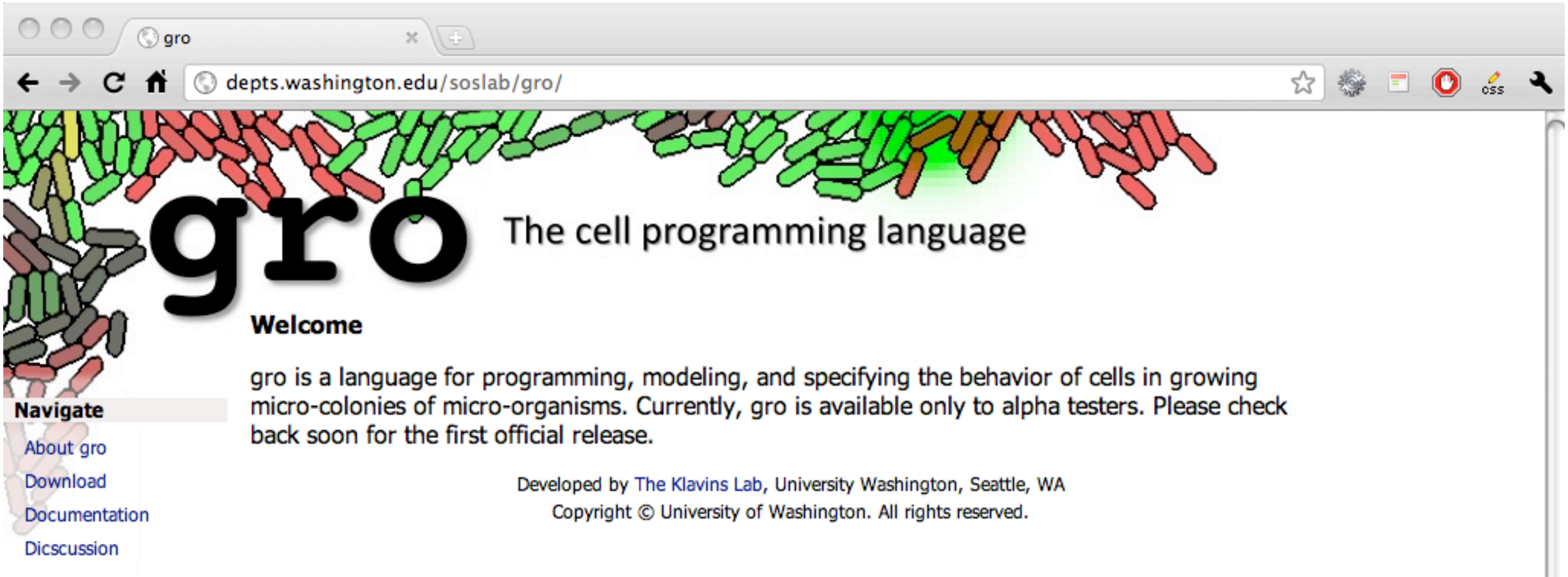


[33] SWAIN P.S., ELOWITZ M.B., SIGGIA E.D.: 'Intrinsic and extrinsic contributions to stochasticity in gene expression', *Science*, 2002, **99**, (20), pp. 12795−12800

Thorsley and Klavins, "Approximating stochastic biochemical processes with Wasserstein pseudometrics", *IET Systems Biology*, June 2009.

Deleted Slides

- Unpublished data and examples deleted.



Send klavins@uw.edu the email address associated with your Google account so that you may download gro!

- Mac OS X 10.5.8 and up
- Windows 7 + Cygwin

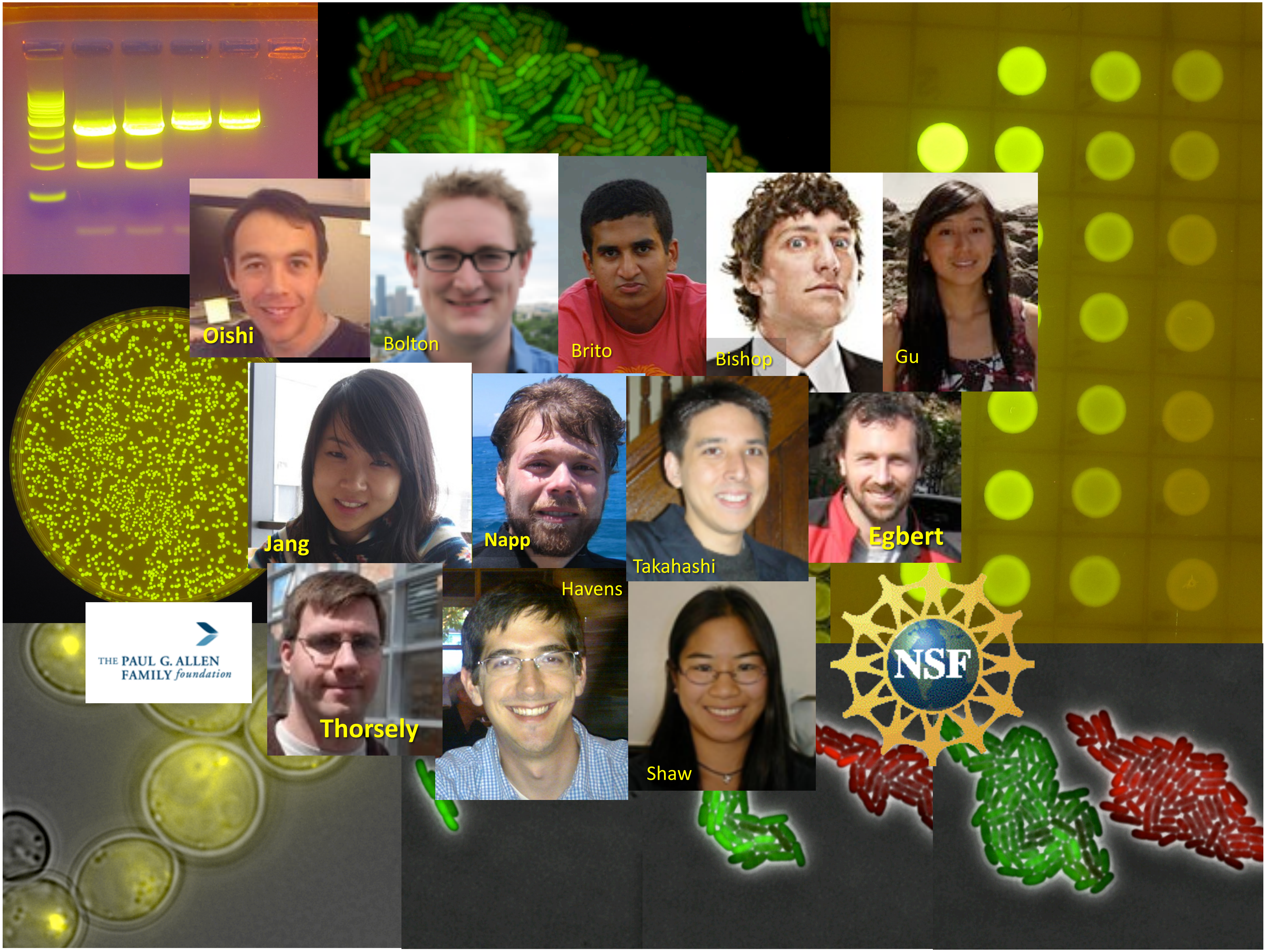
Outline

- Example Experimental Systems
- What Stochasticity Can Do
- Analytical Approaches
  - The Master Equation
  - Moment Dynamics
- Simulation Based Approaches
  - Simulation Methods
  - Approximate Abstraction/Refinement
- Done!

Running Example:
Control of Gene Expression

Example:
Arithmetic

Software
Tools



Oishi
Bolton
Brito
Bishop
Gu
Jang
Napp
Takahashi
Egbert
Havens
THE PAUL G. ALLEN FAMILY foundation
Thorsely
Shaw
NSF