

# Stochastic dynamics of small gene regulation networks

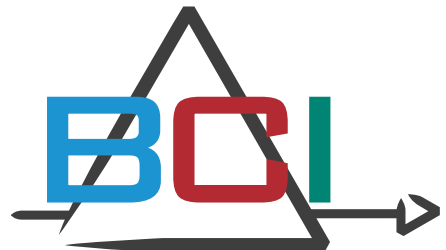
**Lev Tsimring**

*BioCircuits Institute*

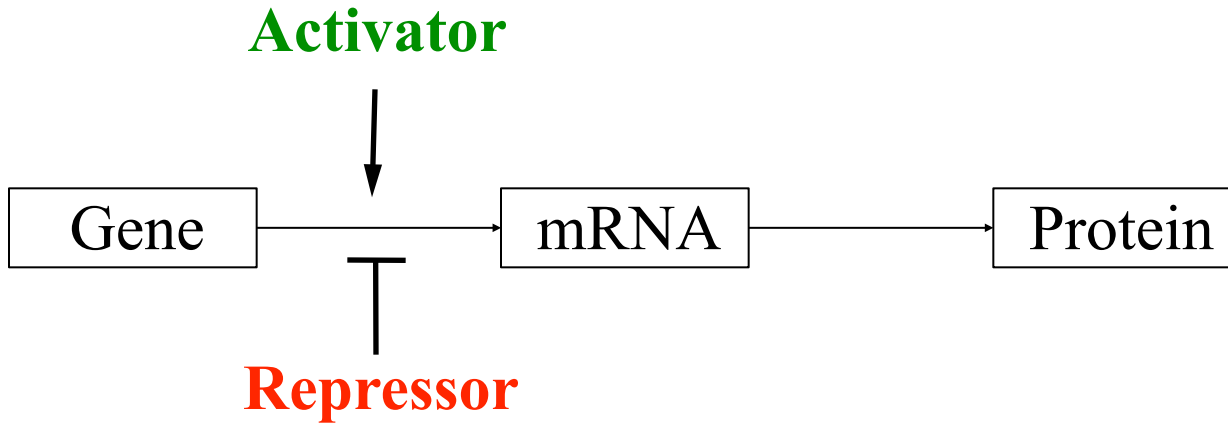
*University of California, San Diego*



*Nizhni Novgorod, June, 2011*



# Central dogma



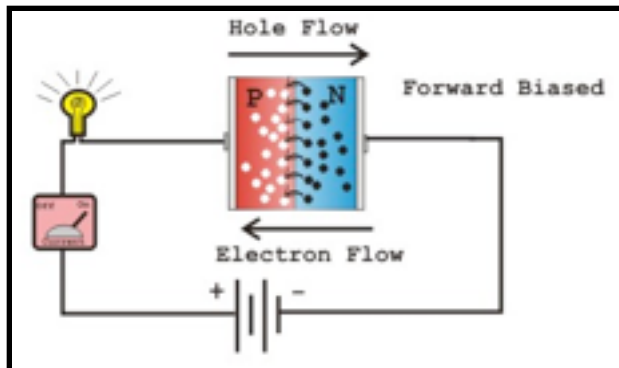
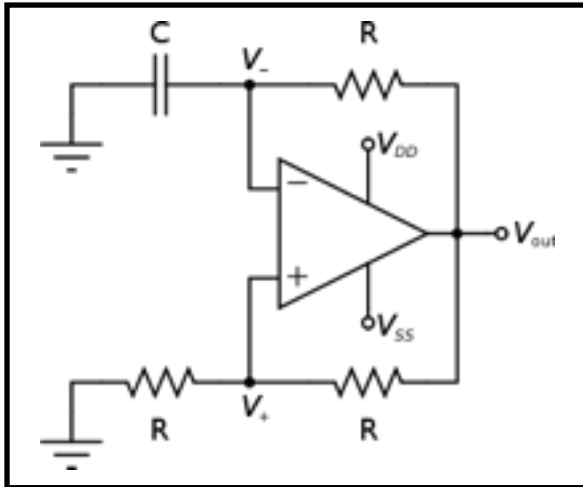
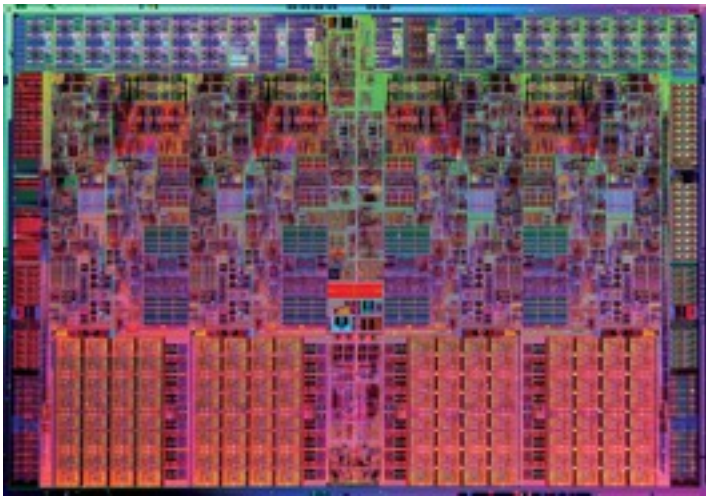
All cells have the same genes...  
...So why cell differ from each other?

## Transcriptional Regulation

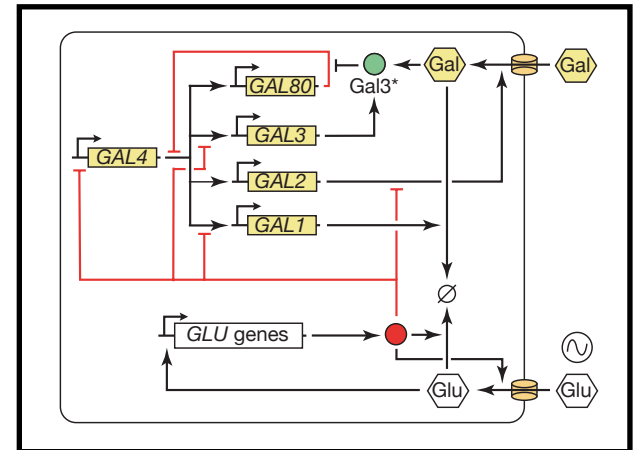
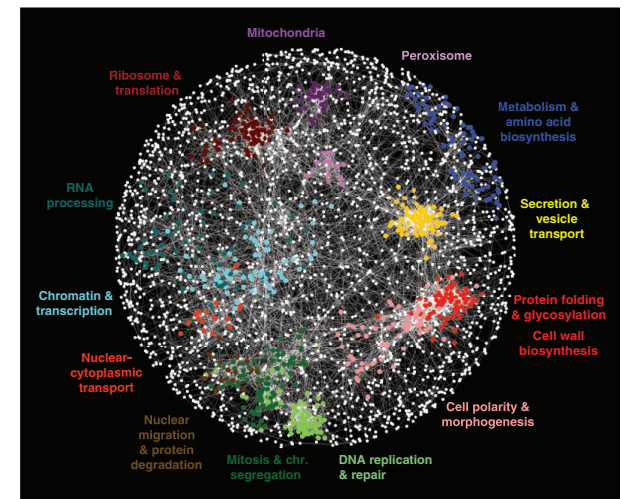
**Activator** – increases rate of production

**Repressor** – decreases rate of production

# Systems Biology

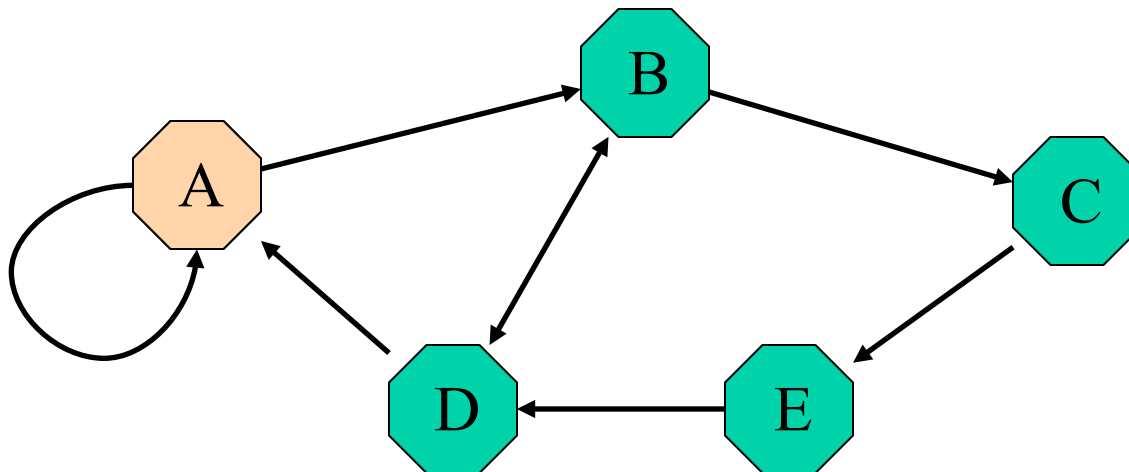


scale

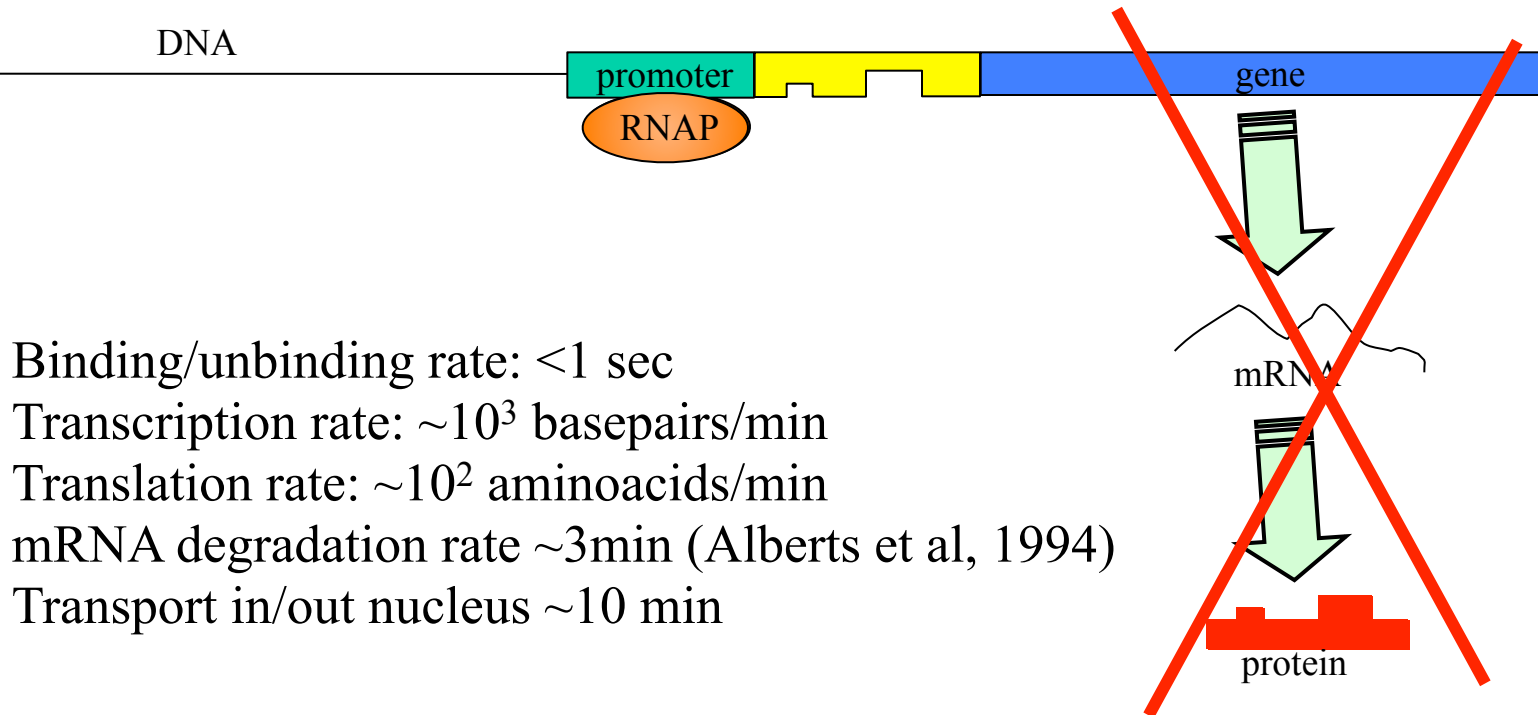


# Gene regulation networks

- Proteins affect rates of production of other proteins (or themselves)
- This leads to formation of *networks* of interacting genes/proteins
- Large stochastic fluctuations
- Sub-networks are *non-Markovian*, even if the whole system is
- Different reaction channels operate at vastly different time scales and number densities



# Auto-repressor: A cartoon

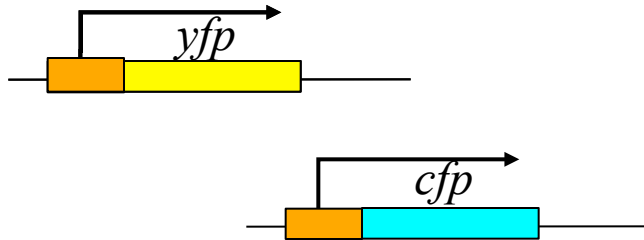


Vastly different time scales: need for reduced descriptions

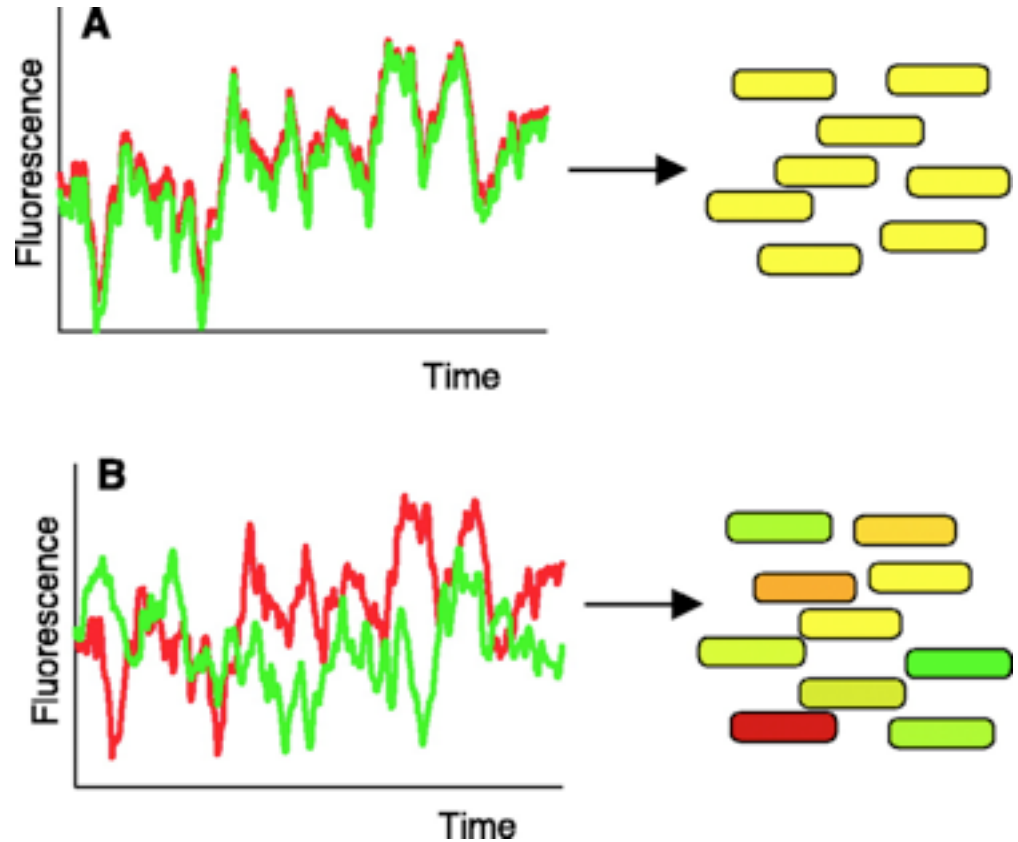
# Transients in gene regulation

- Genetic circuits are never at a fixed point:
  - Intrinsic noise
  - Extrinsic noise
  - External signaling
  - Oscillations:  
circadian rhythms; ultradian rhythms; cell cycle

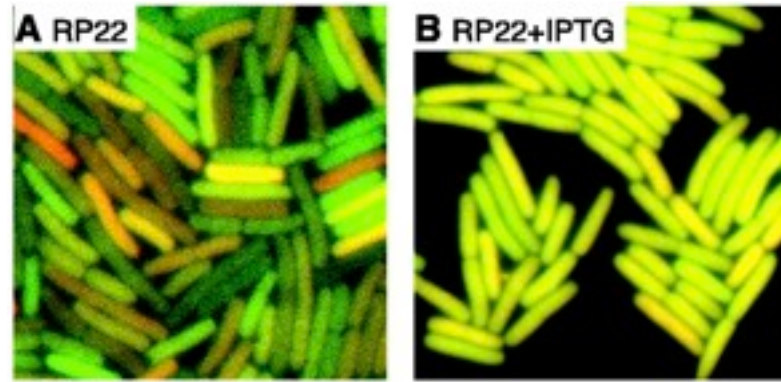
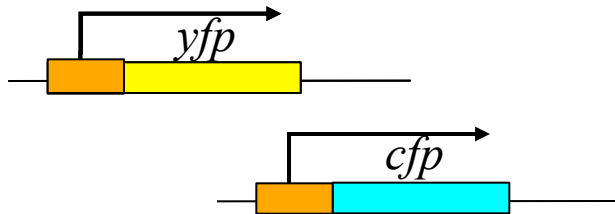
# Intrinsic vs. Extrinsic noise?



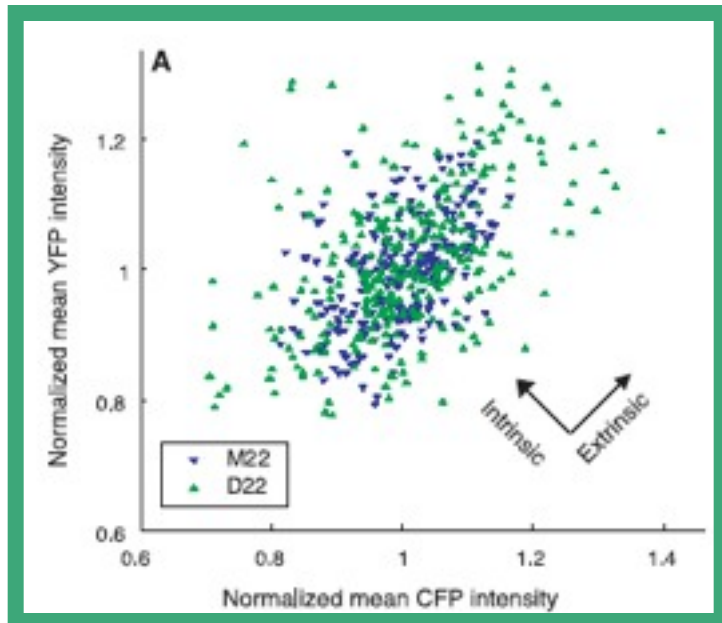
Identical promoters;  
different reporter genes



# Intrinsic vs. Extrinsic noise?

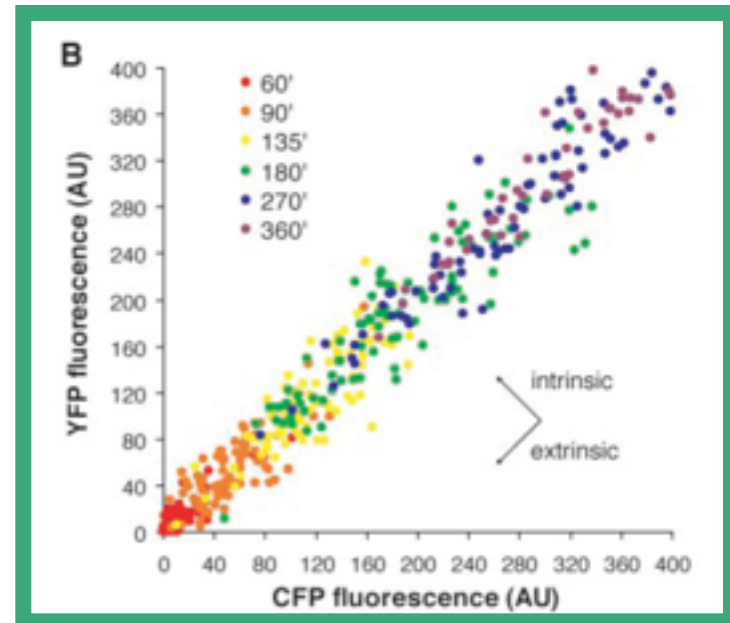


*Bacteria*



Elowitz et al, 2002

*Yeast*

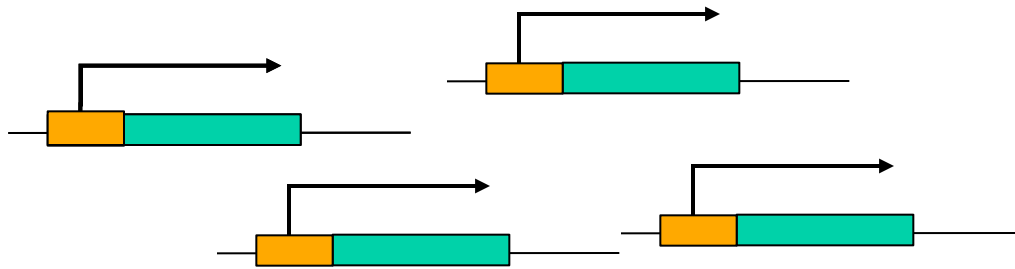


Raser & O'Shea, 2004

Intrinsic noise is smaller in Yeast



# Intrinsic vs. Extrinsic noise



Multiple identical promoters;  
same reporter gene

Gene expression:

$$G_M = \sum_{i=1}^M \langle g_i \rangle; \quad V_M = \sum_{i=1}^M \left[ \langle \tilde{g}_i^2 \rangle + \sum_{j \neq i} \langle \tilde{g}_i \tilde{g}_j \rangle \right]$$

For identical promoter-gene pairs:

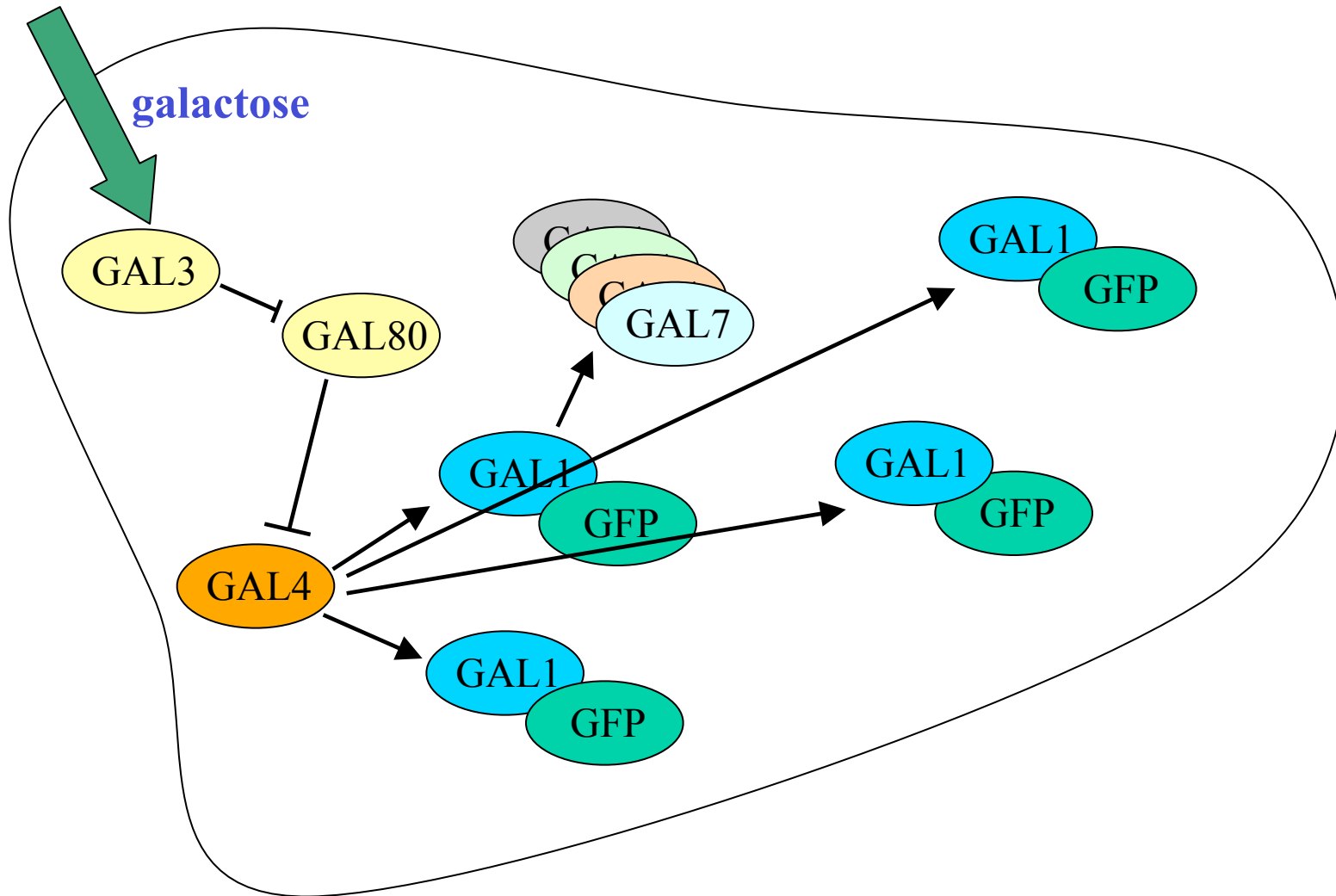
$$G_M = Mg; \quad V_M = Mv + M(M-1)c$$

For intrinsic noise ( $c=0$ ),  $V_M = Mv$ ;  $CV \propto M^{-1/2}$

For correlated (extrinsic) noise ( $c=v$ ),  $V_M = M^2v$ ;  $CV = \text{const}(M)$

# Intrinsic vs. Extrinsic noise

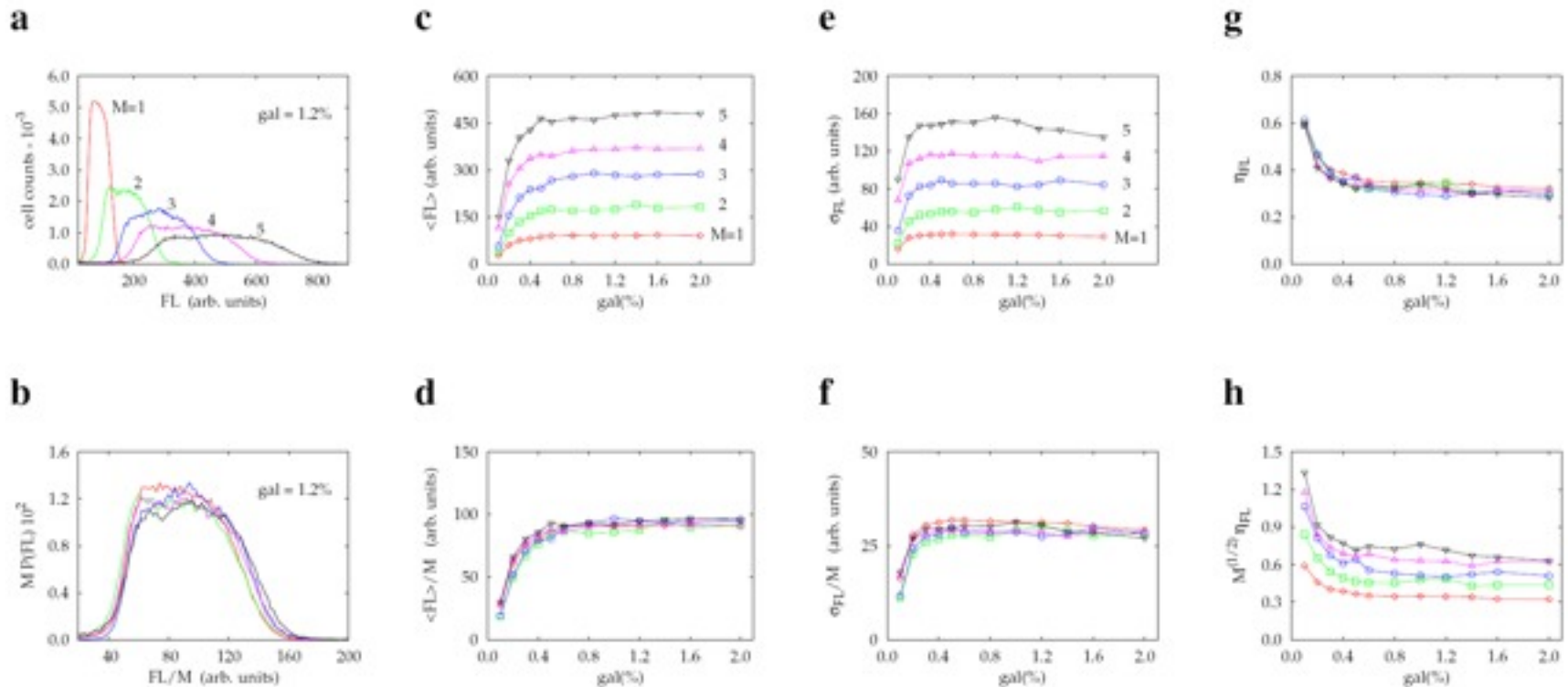
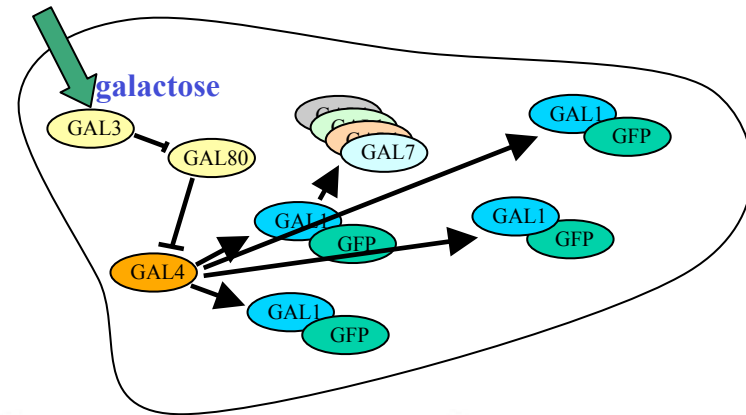
*Saccharomyces cerevisiae* (yeast); galactose utilization circuit



# Intrinsic vs. Extrinsic noise

Volfson et al., Nature, 2006

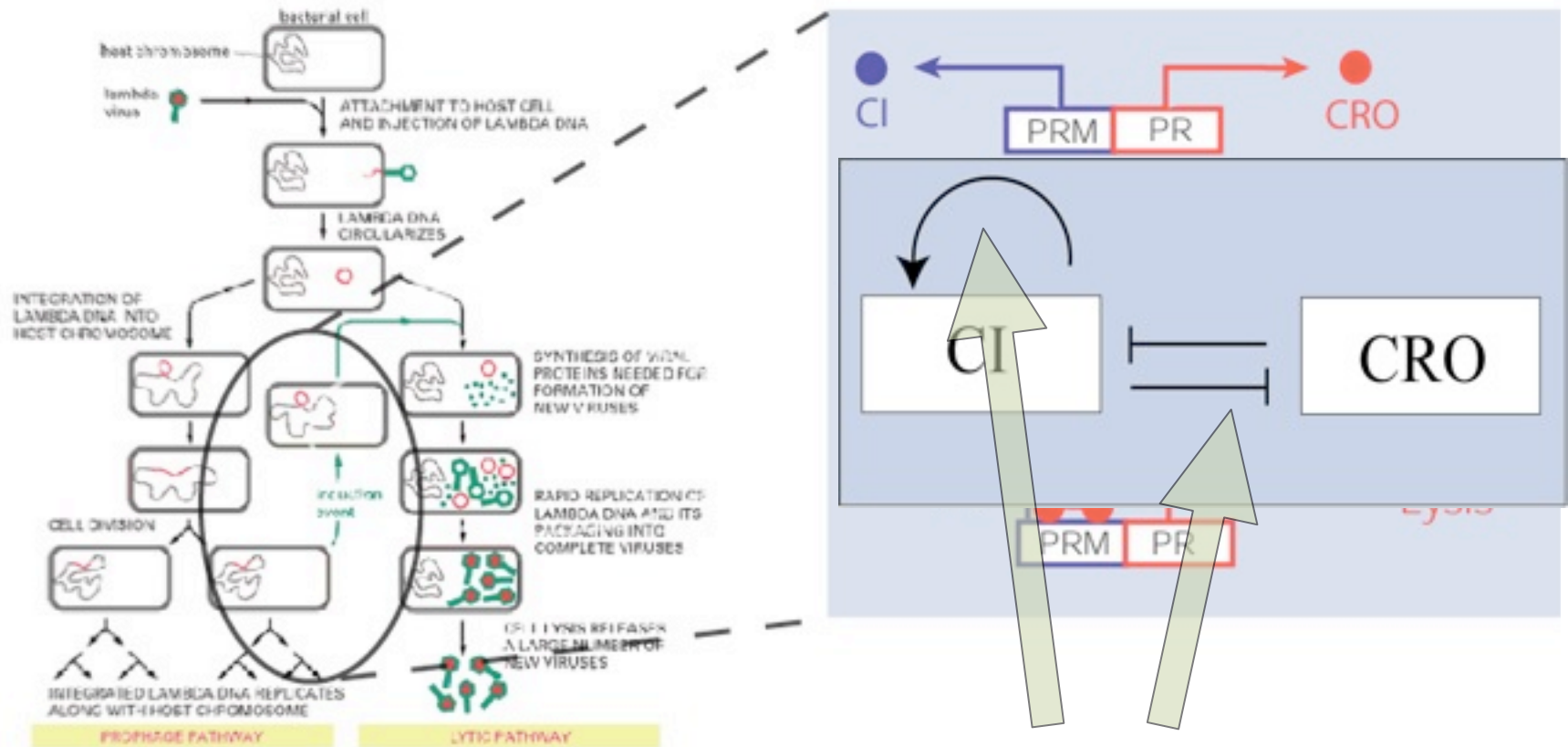
Flow cytometry data



Intrinsic noise is small in Yeast

# Bistability: $\lambda$ -Phage Life Cycle

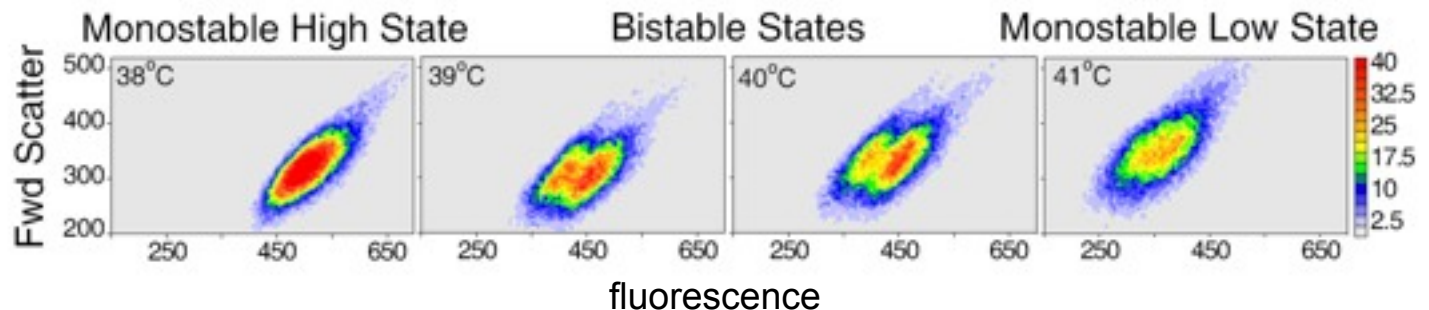
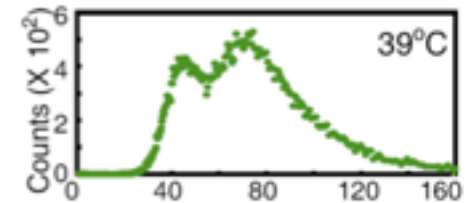
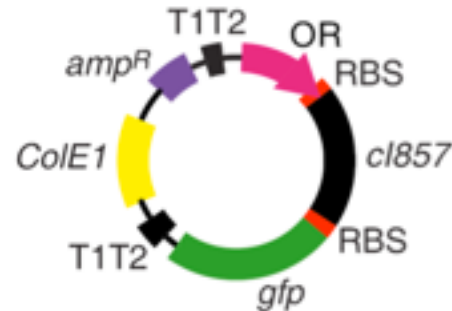
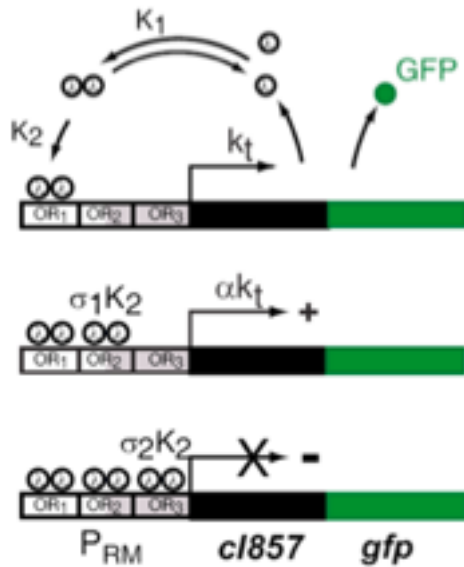
M.Ptashne, 2002



Two positive feedback loops

# Genetic noise

## Experiments with synthetic bistable autoregulatory circuit in *E. coli* (derived from $\lambda$ -phage)

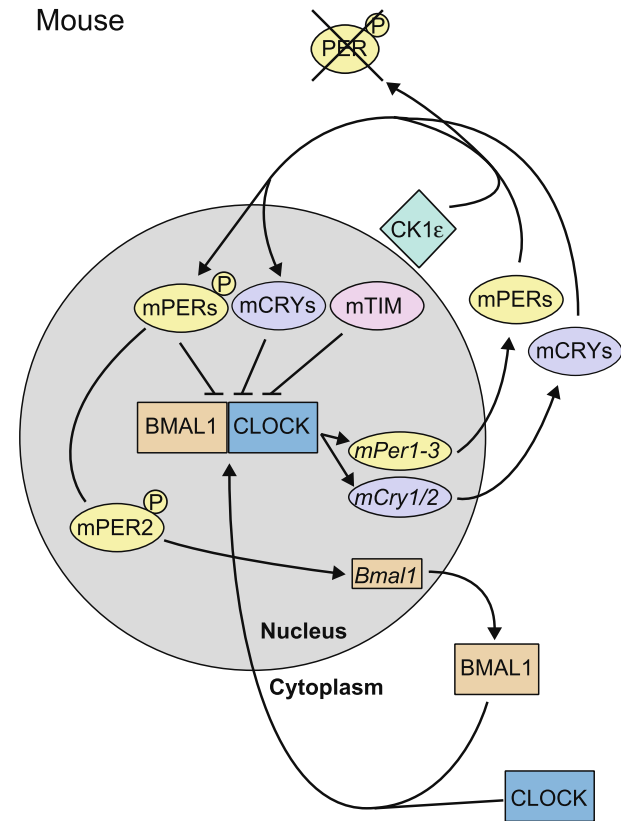
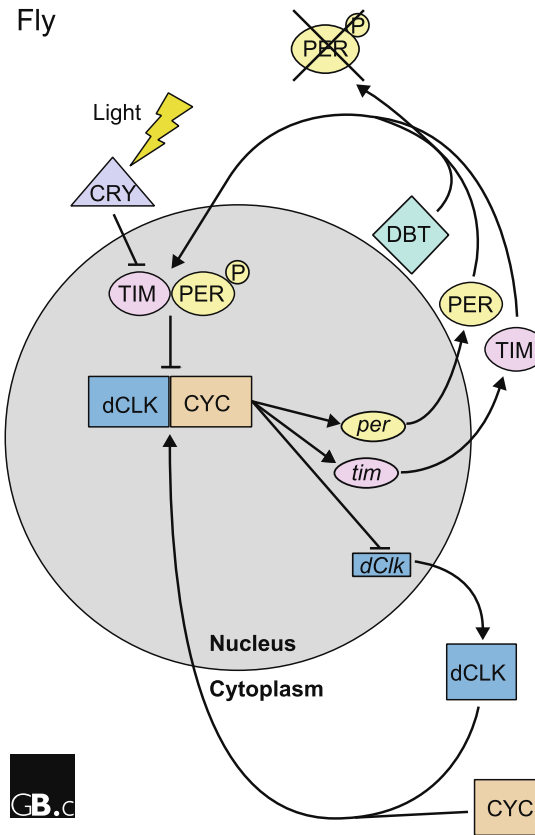


# Oscillations: circadian clock

*Neurospora crassa*



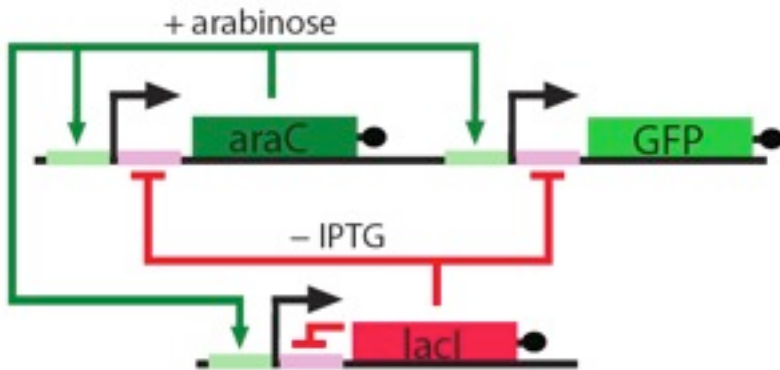
Courtesy of Stu Brody



# Synthetic gene oscillator in *E.coli*

Stricker *et al.*, Nature, 456(7221):516-9 (2008).

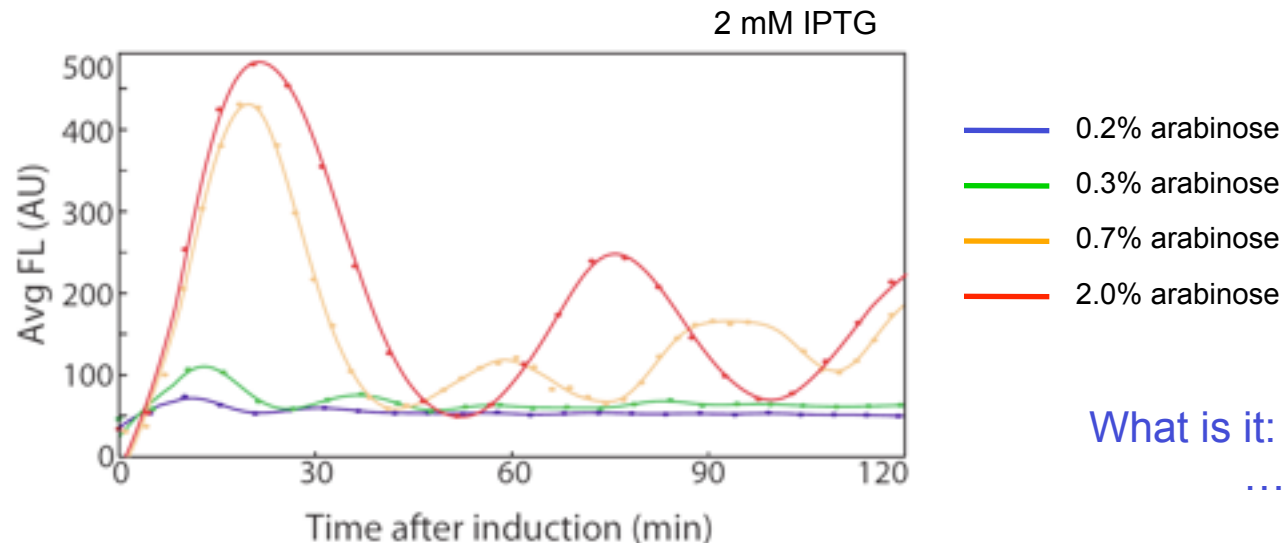
Design: use AraC as activator and LacI as repressor



- Three plasmid with identical promoters:  
(i) activator, (ii) repressor, (iii) reporter
- All proteins are tagged with ssRA tag for fast enzymatic decay (~15-20min)
- Two experimental “knobs”: Arabinose required for activator binding and IPTG prohibits repressor binding

Flow cytometry:

time series of fluorescence in initially synchronized batch culture of cells

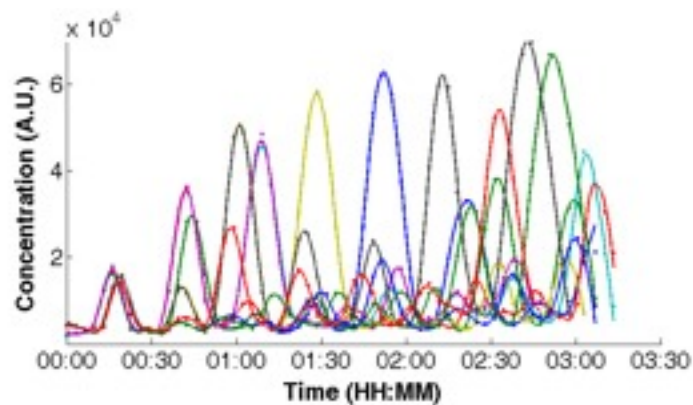
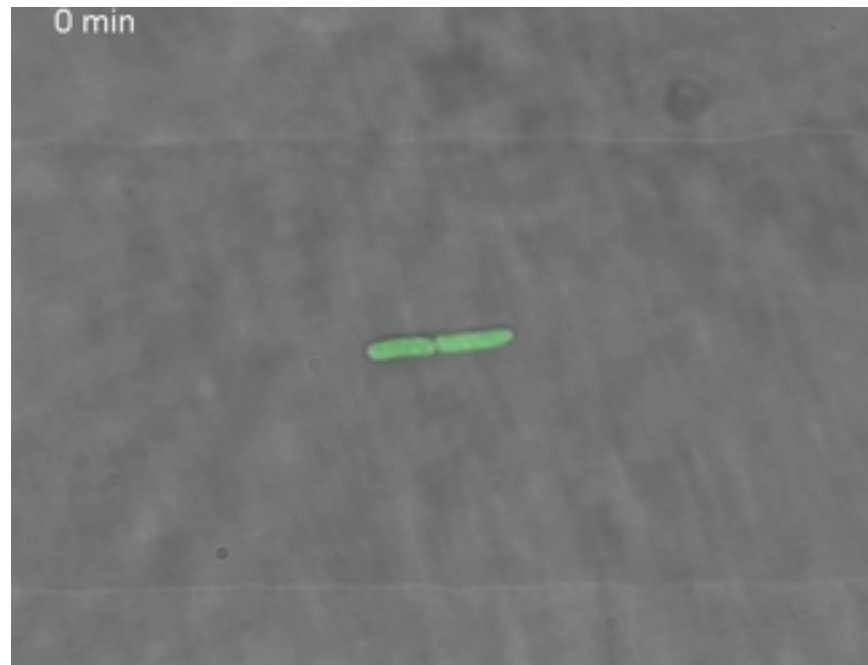


What is it: oscillations stop  
...or desynchronize?

# Single cell experiments

Stricker *et al.*, Nature, 456(7221):516-9 (2008).

2mM IPTG  
0.7% arabinose



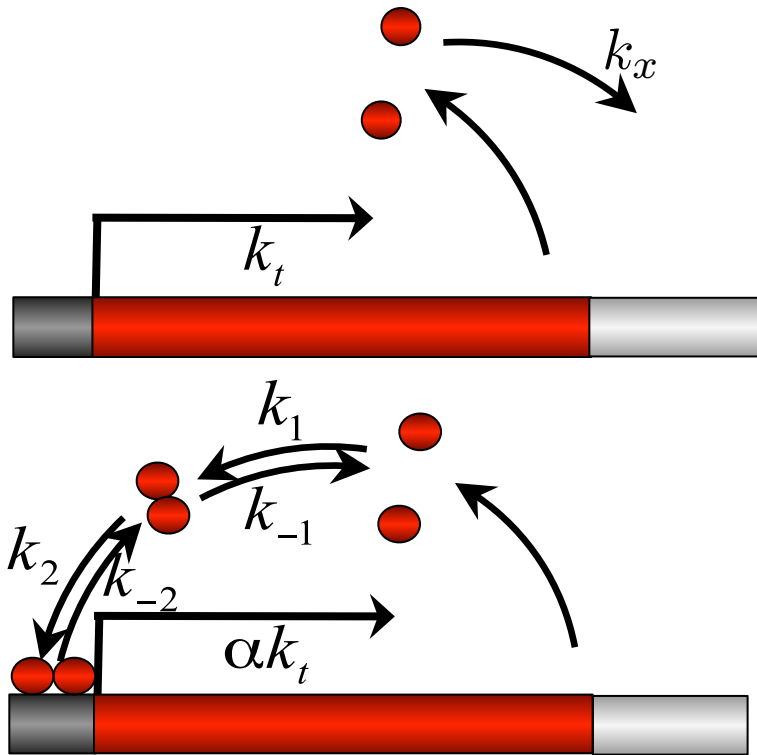


# Modeling gene regulation

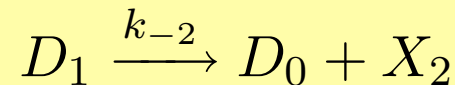
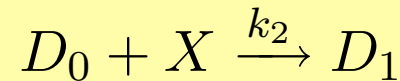
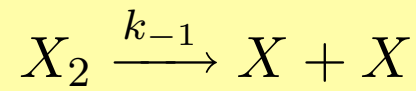
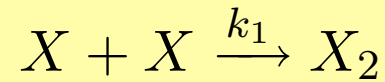
Techniques for modeling gene networks dynamics

- Boolean dynamics
- Deterministic Equations (Mass Action kinetics)
- Monte-Carlo Simulations; Master Equations
- Stochastic Differential Equations

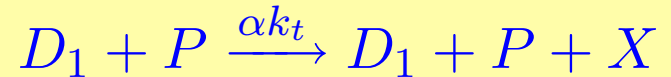
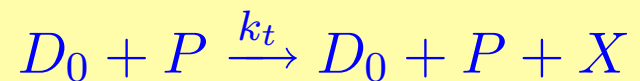
# Modeling gene regulation: Single gene circuit



Fast



Slow



Binding/unbinding rate:  $\leq 1$  sec

Transcription rate:  $\sim 10^3$  basepairs/min

Translation rate:  $\sim 10^2$  aminoacids/min

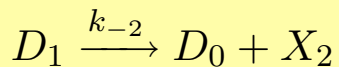
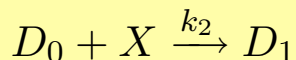
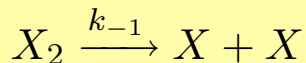
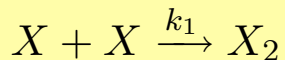
mRNA degradation rate  $\sim 3$ min

Transport in/out of nucleus 10min..10hrs

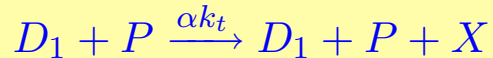
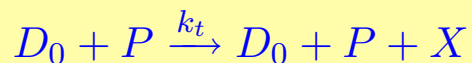
Protein degradation 10min..hours

# Mass action kinetics

Fast



Slow



$$\dot{x} = -2k_1x^2 + 2k_{-1}x_2 - k_x x + k_t p_0(d_0 + \alpha d_1)$$

$$\dot{x}_2 = k_1x^2 - k_{-1}x_2 - k_2d_0x_2 + k_{-2}d_1$$

$$\dot{d}_0 = -k_2d_0x_2 + k_{-2}d_1$$

$$\dot{d}_1 = k_2d_0x_2 - k_{-2}d_1$$

Separation of scales:

$$x_t = x + 2x_2 + 2d_1$$

$$d = d_0 + d_1$$

$$x_2 = K_1x^2$$

$$K_1 = k_1/k_{-1}$$

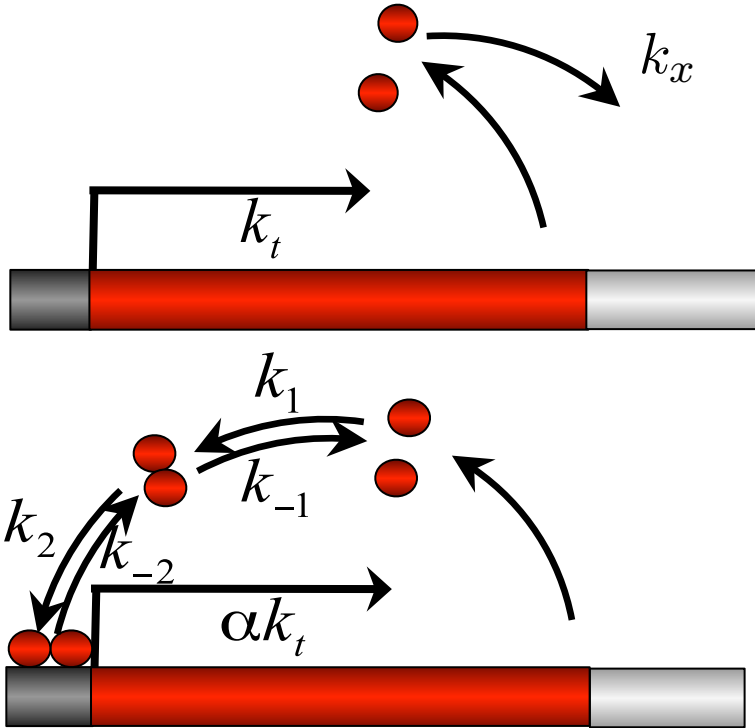
$$d_1 = K_2d_0x_2$$

$$K_2 = k_2/k_{-2}$$

$$\dot{x}_t = \mathcal{P}_x \dot{x} = k_t p_0 d \frac{1 + \alpha K_1 K_2 x^2}{1 + K_1 K_2 x^2} - k_x x$$

$$\mathcal{P}_x = x'_t(x) = 1 + 4K_1x + \frac{4dK_1K_2x^2}{1 + K_1K_2x^2}$$

# Including fluctuations: Monte-Carlo Simulations



Time between reactions:

$$\tau = (1/\sum_i P_i) \ln(1/r)$$

Jump probabilities:

$$P_1(2X \rightarrow X - 2) = k_1 x(x - 1)$$

$$P_2(X_2 \rightarrow 2X) = k_{-1} x_2$$

$$P_3(D_0 \rightarrow D_0 + X) = k_t d_0 x$$

$$P_4(D_1 \rightarrow D_1 + X) = \alpha k_t d_1 x$$

$$P_5(D_0 \rightarrow D_1) = k_2 d_0 x$$

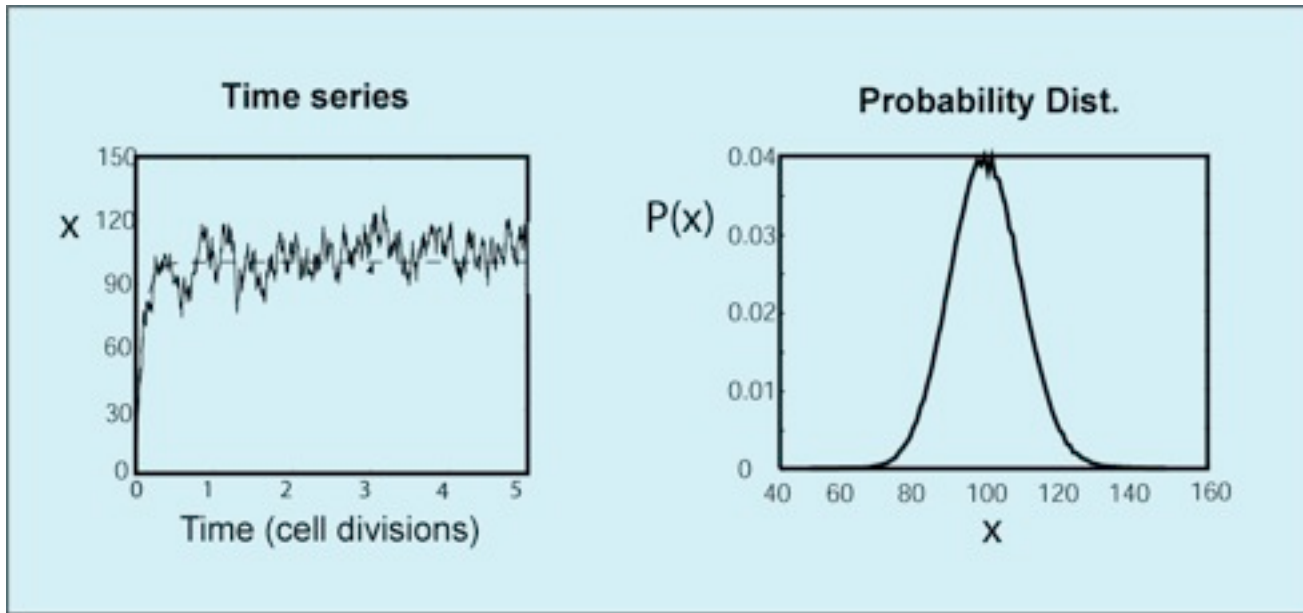
$$P_6(D_1 \rightarrow D_0) = k_{-2} d_1$$

$$P_7(X \rightarrow) = k_x x$$

Simulations:

1. Use  $(x, x_2, d)$  to determine  $P_i$
2. Determine  $\tau$
3. Determine specific reaction and update
4. Loop

# Monte Carlo Data



Complete and exact

Straightforward to simulate

Computationally expensive

Nonintuitive

*Since it is exact, it is the gold standard for evaluating the accuracy of other modeling approaches*

# Master equation description

$n$  – total # of monomers;  $u$  – # of unbound dimers;  $b$  – # of bound dimers

Master equation for  $p_{n,u,b}$

$$\begin{aligned} \frac{dp_{n,u,b}}{dt} = & k_x [(n+1-2u-2b)p_{n+1,u,b} - (n-2u-2b)p_{n,u,b}] \\ & + k_t (d-b+\alpha b)[p_{n-1,u,b} - p_{n,u,b}] \\ & + k_1 [(n-2u-2b+2)(n-2u-2b+1)p_{n,u-1,b} \\ & - (n-2u-2b)(n-2u-2b-1)p_{n,u,b}] \\ & - k_{-1} [(u+1)p_{n,u+1,b} - dp_{n,u,b}] \\ & + k_{-2} [(b+1)p_{n,u-1,b+1} - bp_{n,u,b}] - k_2 [u(d-b)p_{n,u,b} - (u+1)(d-b+1)p_{n,u+1,b-1}] \end{aligned}$$

Projection:  $p_{n,u,b} = p_n p_{u,b|n}$

$$\begin{aligned} \frac{dp_n}{dt} = & k_x [(n+1-2\langle u|n+1\rangle - 2\langle b|n+1\rangle)p_{n+1} - (n-2\langle u|n\rangle - 2\langle b|n\rangle)p_n] \\ & + k_t [(d+(\alpha-1)\langle b|n-1\rangle)p_{n-1} - (d+(\alpha-1)\langle b|n\rangle)p_n] \end{aligned}$$

$$p_{u,b|n} = p_0 \frac{K_1^{u+b} K_2^b d! n!}{b!(d-b)! u! (n-2u-2b)!}$$

$$\langle u|n \rangle = K_1 m(m-1); \quad \langle b|n \rangle = \frac{dK_1 K_2 m(m-1)}{1 + K_1 K_2 m(m-1)}$$

$$m = n - 2\langle u|n \rangle - 2\langle b|n \rangle$$

See Kepler & Elston, Biophys. J., 2001

# Back to ODE

In the continuum limit (large  $n$ ): Fokker-Planck equation

$$\frac{\partial \rho(x_t)}{\partial t} = \frac{\partial}{\partial x_t} \left[ \left( k_x x - k_t d \frac{1 + \alpha K_1 K_2 x^2}{1 + K_1 K_2 x^2} \right) \rho \right] + \frac{1}{2} \partial_{x_t}^2 \left[ \left( k_x x + k_t d \frac{1 + \alpha K_1 K_2 x^2}{1 + K_1 K_2 x^2} \right) \rho \right]$$

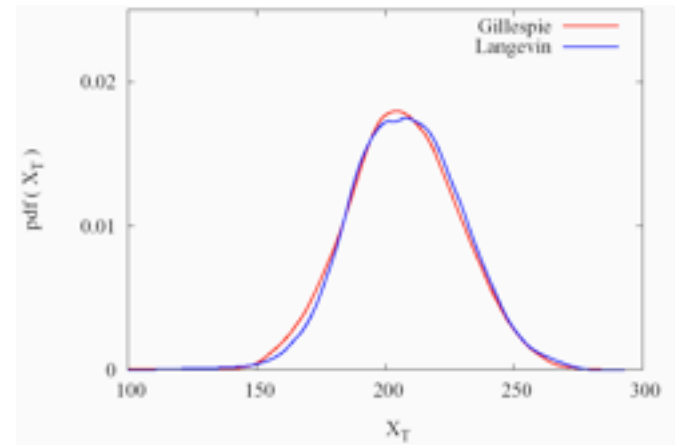
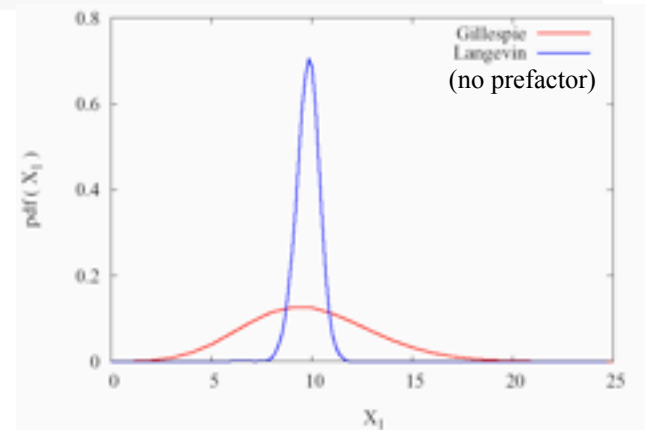
$$x_t = x + 2K_1 x^2 + \frac{2dK_1 K_2 x^2}{1 + K_1 K_2 x^2}$$

Corresponding Langevin equation

$$\frac{dx_t}{dt} = k_t d \frac{1 + \alpha K_1 K_2 x^2}{1 + K_1 K_2 x^2} - k_x x + \sqrt{D(x_t)} \xi$$

$$D(x) = k_x x + k_t d \frac{1 + \alpha K_1 K_2 x^2}{1 + K_1 K_2 x^2}$$

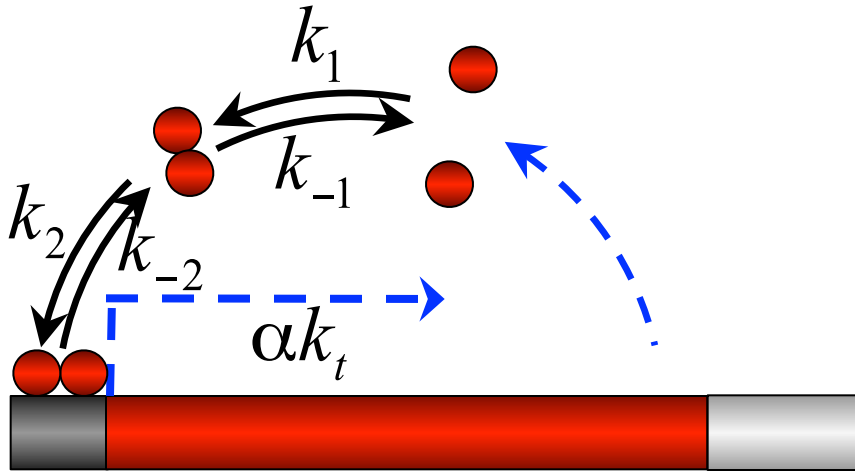
with  $x = x(x_t)$



Fast reaction noise is filtered out

Kepler & Elston, *Biophys J* **81**:3116 (2001)

# Transcriptional delay



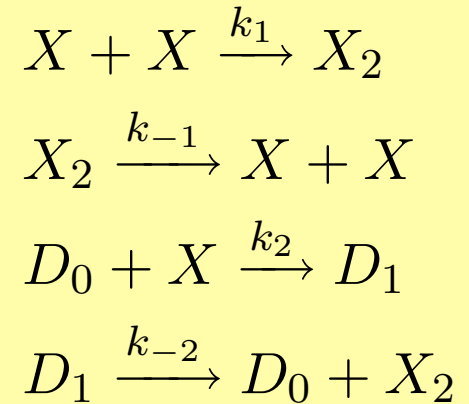
$$\begin{aligned}\dot{x} &= -2k_1x^2 + 2k_{-1}x_2 - k_x x + k_t p_0(d_0(t-T) + \alpha d_1(t-T)) \\ \dot{d}_0 &= -k_2d_0x_2 + k_{-2}d_1 \\ \dot{d}_1 &= k_2d_0x_2 - k_{-2}d_1 \\ \dot{x}_2 &= k_1x^2 - k_{-1}x_2 - k_2d_0x_2 + k_{-2}d_1\end{aligned}$$

[cf. Santillán & Mackey, 2001]

After projection

$$\mathcal{P}_x \dot{x} = k_t p_0 d \frac{1 + \alpha K_1 K_2 x^2 (t-T)}{1 + K_1 K_2 x^2 (t-T)} - k_x x$$

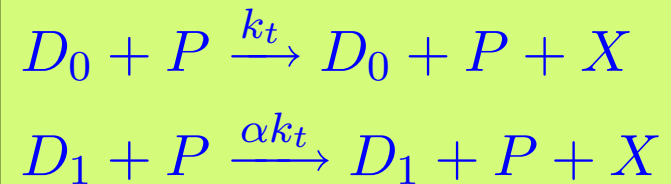
Fast



Slow



Delayed





# Genetic oscillations: Hopf bifurcation

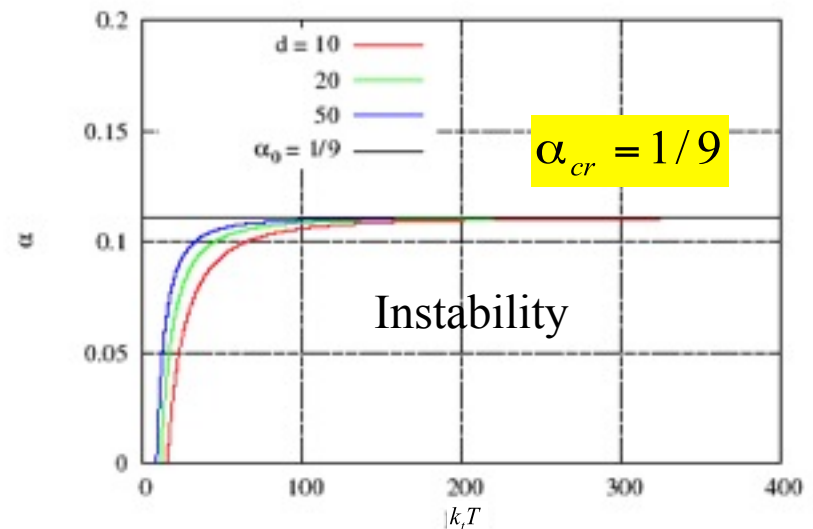
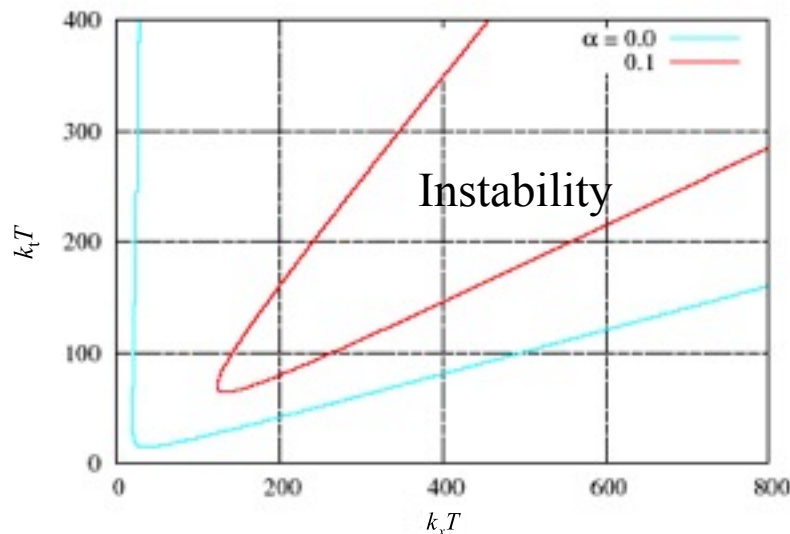
Fixed point:

$$0 = -k_x \bar{x} + k_t d H(\bar{x}) \quad H(\bar{x}) = \frac{2K_1 K_2 (\alpha - 1)}{(1 + K_1 K_2 \bar{x}^2)^2}$$

Complex eigenvalues

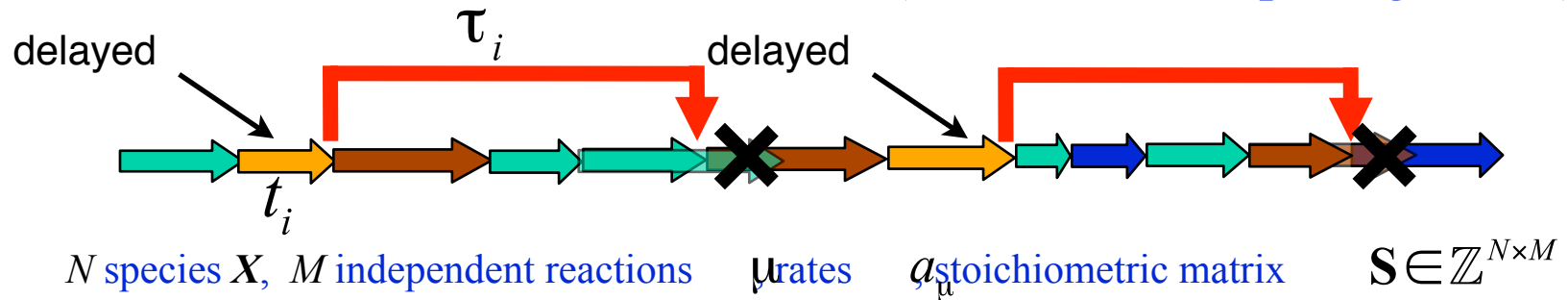
$$\mathcal{P}_x(\bar{x})\lambda = -k_x + k_t d G(\bar{x}) e^{-\lambda T} \quad G(\bar{x}) = \frac{dH(\bar{x})}{d\bar{x}}$$

$$d = 10, K_1 = K_2 = 1$$



# Stochastic simulations of delayed reactions

(modified Gillespie algorithm)



## Markovian statistics of reaction times:

- exponential “next reaction” time distribution

$$P(t) \propto \exp \left[ -t \sum_{\mu} a_{\mu} \right]$$

- which reaction to choose?

$$P(\mu = \mu') = a_{\mu'} / \sum_{\mu} a_{\mu}$$

- immediate execution



## Delayed reactions [e.g. transcription, translation]:

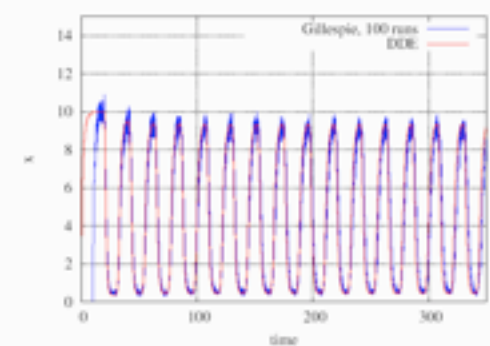
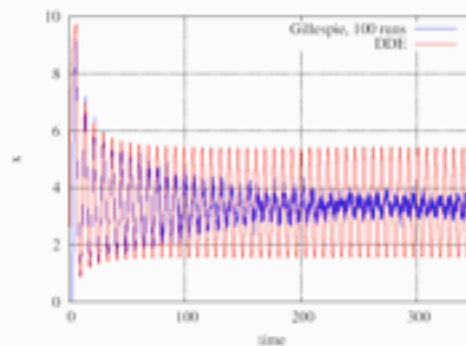
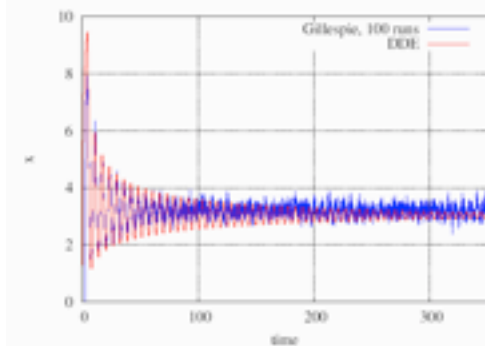
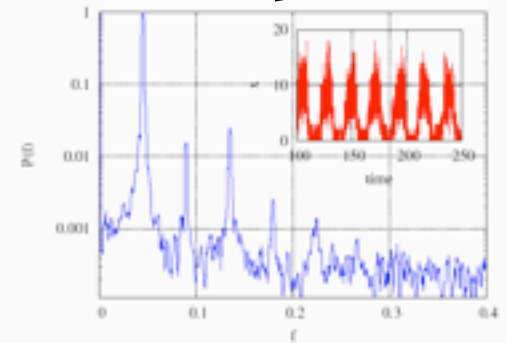
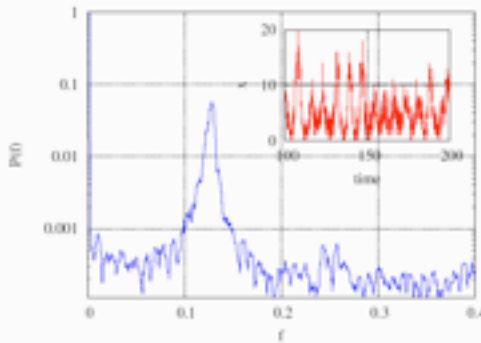
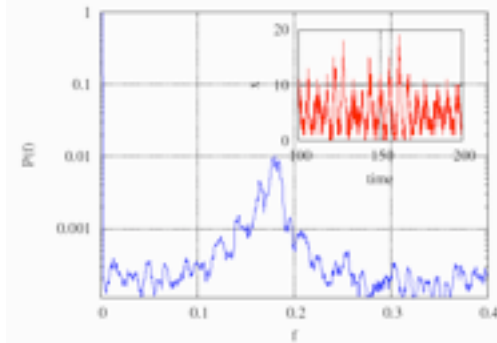
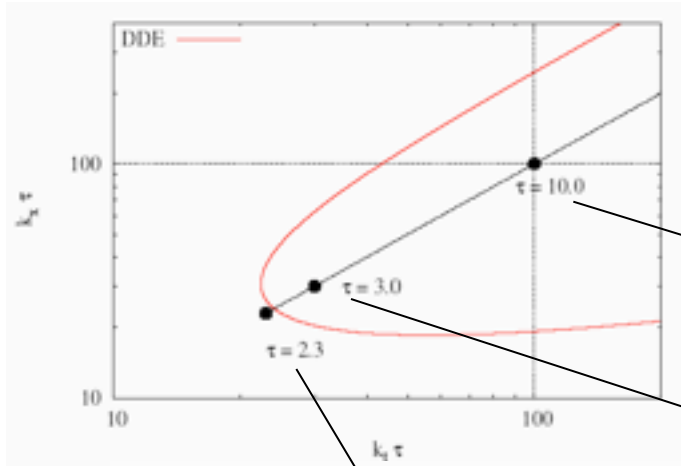
- Delayed execution
- Execution time  $\tau$  : fixed or stochastic,  
e.g. with Gaussian time distribution

$$P_{\mu}(\tau) \propto \exp \left[ -(\tau - \tau_0)^2 / \sigma_{\mu}^2 \right]$$

# Stochastic simulations

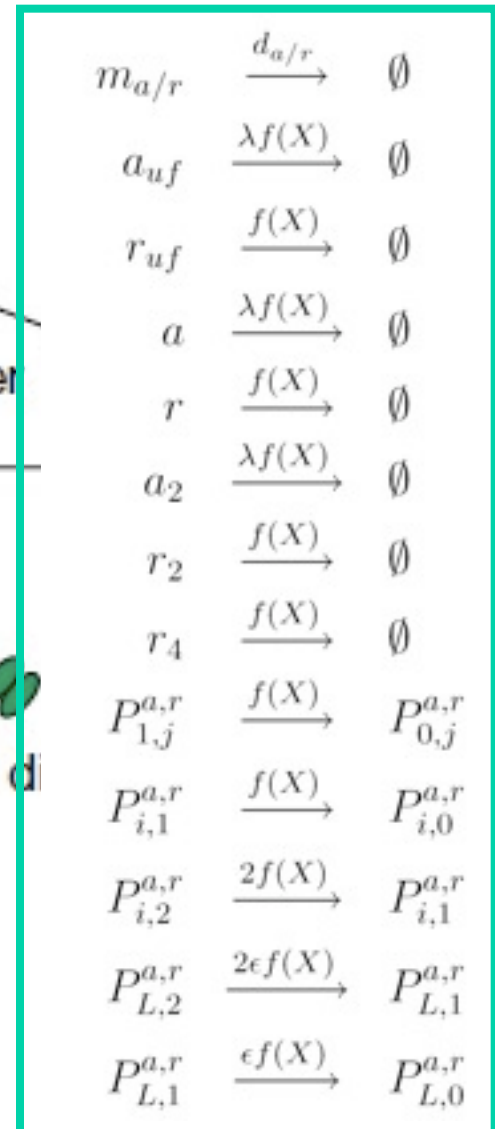
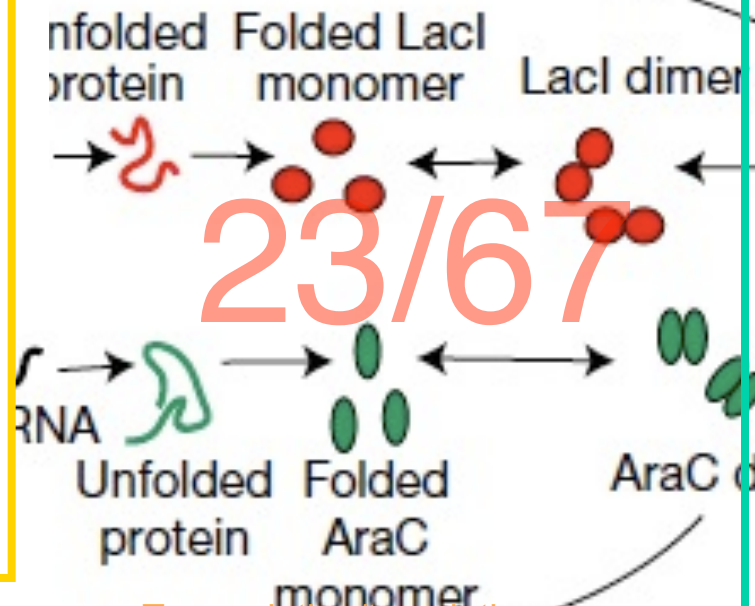
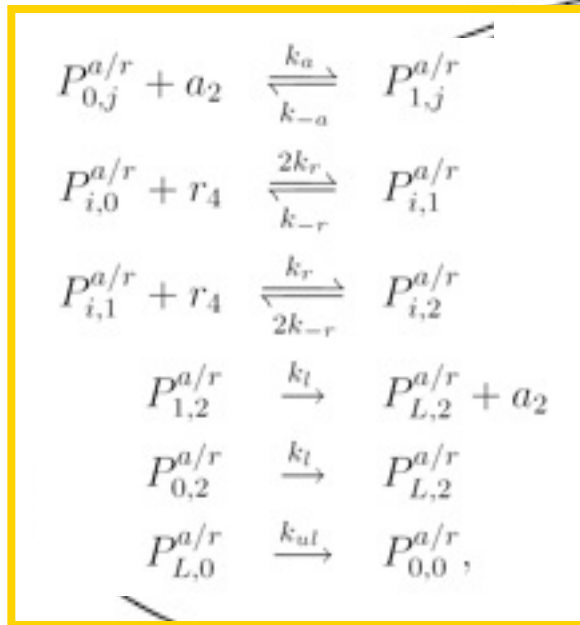
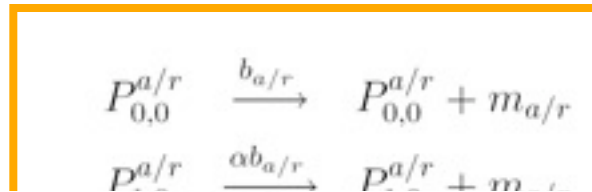
Bratsun, Volfson, LT, Hasty, *PNAS*, 2005

$$d = 10, k_1 = k_2 = 100, K_1 = K_2 = 0.5, \alpha = 0$$



# Synthetic bacterial oscillator: A detailed model

Stricker et al., Nature, 456(7221):516-9 (2008).



Regulation

Transcription/translation

$$k_r = k_{-r} \left( (C_r^{max} - C_r^{min}) \frac{1}{1 + \left( \frac{[IPTG]}{k_{r1}} \right)^{b_1}} + C_r^{min} \right)$$

$$k_a = k_{-a} \left( (C_a^{max} - C_a^{min}) \cdot \frac{[ara]^{c_1}}{k_{a1}^{c_1} + [ara]^{c_1}} \cdot \frac{1}{1 + \left( \frac{[IPTG]}{k_{r2}} \right)^{b_2}} + C_a^{min} \right)$$

$\begin{aligned}
 i &\in \{0,1\} \\
 j &\in \{0,1,2\}
 \end{aligned}$

Enzymatic decay

$$f(X) = \frac{\gamma}{c_e + X}$$

# Explicit delay model

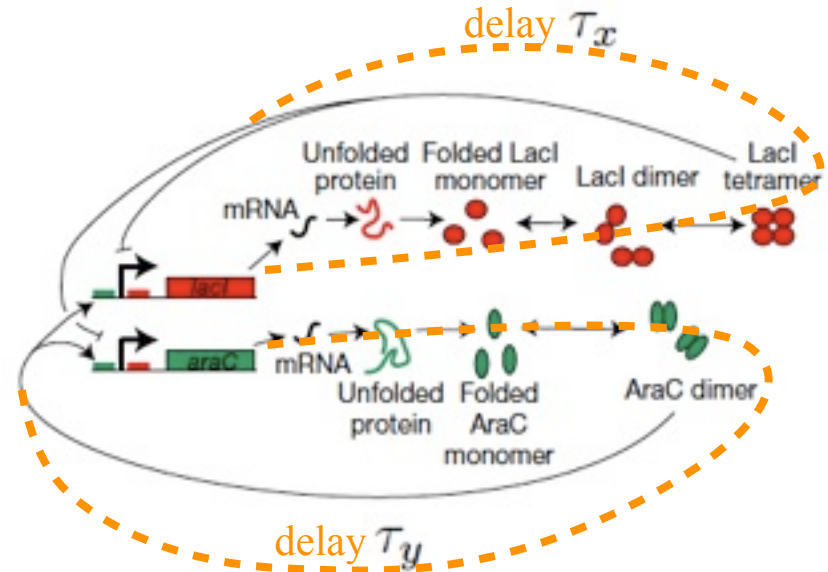
replace chains of interactions with a single effective delay in feedback

lumps many potentially unknown parameters into a few meaningful parameters (identify ignorance)

makes mechanism more transparent

amendable to analytic treatment

faster simulations using delay approximation



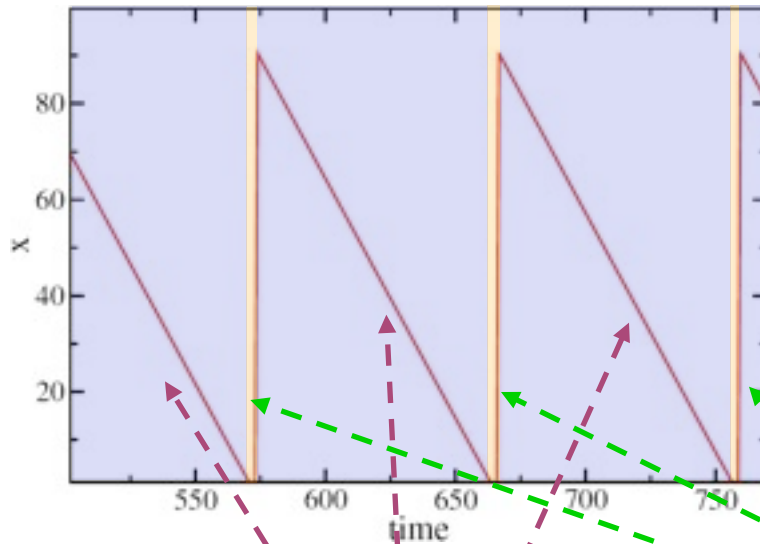
$$\frac{dx}{dt} = \frac{\alpha(1 + fy_{\tau_y}^2/C_{0y}^2)}{(1 + x_{\tau_x}^2/C_{0x}^2)(1 + y_{\tau_y}^2/C_y^2)} - \frac{\delta_x x}{C_{1x} + x}$$

$$\frac{dy}{dt} = \frac{\alpha(1 + fy_{\tau_y}^2/C_{0y}^2)}{(1 + x_{\tau_x}^2/C_{0x}^2)(1 + y_{\tau_y}^2/C_y^2)} - \frac{\delta_y y}{C_{1y} + y}$$

# Delayed auto-repression: degrade-and-fire model

Mather et al, PRL 102, 068105 (2009)

$$\tau = 1; \alpha = 100; C_0 = C_1 = 0.01; \delta = 1$$



$$\dot{x}(t) = \frac{\alpha}{\left(1 + \frac{x(t-\tau)}{C_0}\right)^2} - \delta \frac{x(t)}{C_1 + x(t)}$$

$\alpha$  large;  $C_0, C_1$  small

Fast production

$$\dot{x}_p(t) = \frac{\alpha}{\left(1 + \frac{x_p(t-\tau)}{C_0}\right)^2}$$

Slow degradation

$$\dot{x}_d(t) = -\delta \frac{x_d(t)}{C_1 + x_d(t)}$$

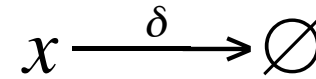
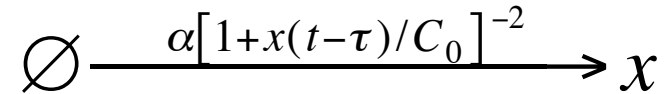
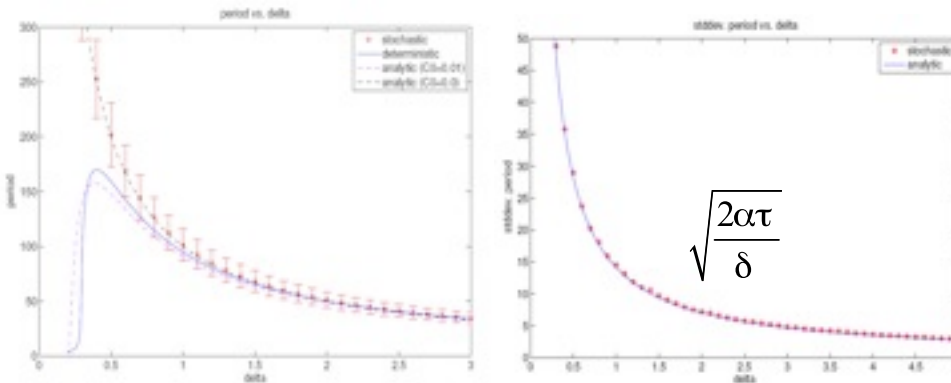
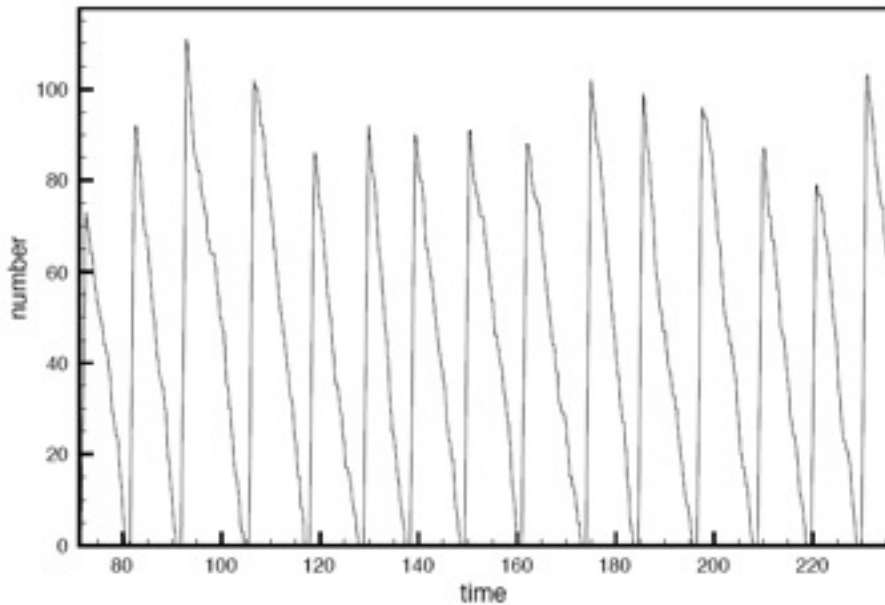
- Two time scales: relaxation oscillations

Period:

$$T \approx \alpha t_c / \delta + \tau + t_c \approx \alpha \tau / \delta$$

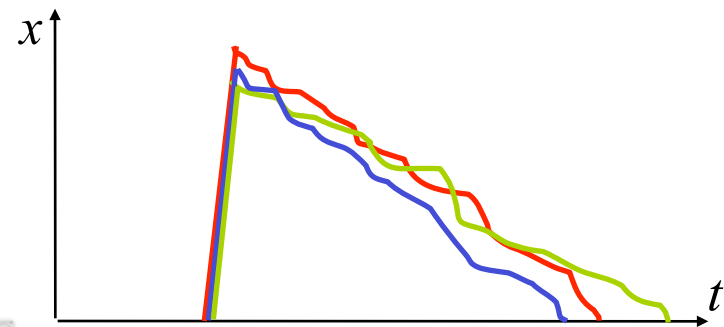
# Zeroth-order degradation: stochastic model

$$\tau = 1; \alpha = 100; C_0 = 0.01; \delta = 1$$



For  $C_0 \rightarrow 0$  these two reactions do not overlap in time

Period variability:



Two sources of variability (both Poissonian):

- fluctuations of  $x_{\max}$

$$v_x = \langle x_{\max} \rangle = \alpha\tau$$

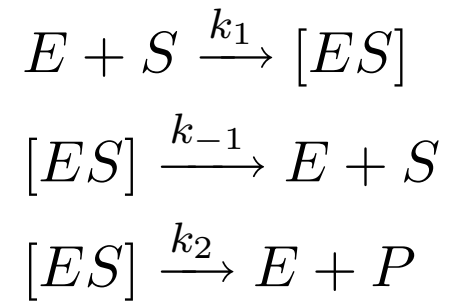
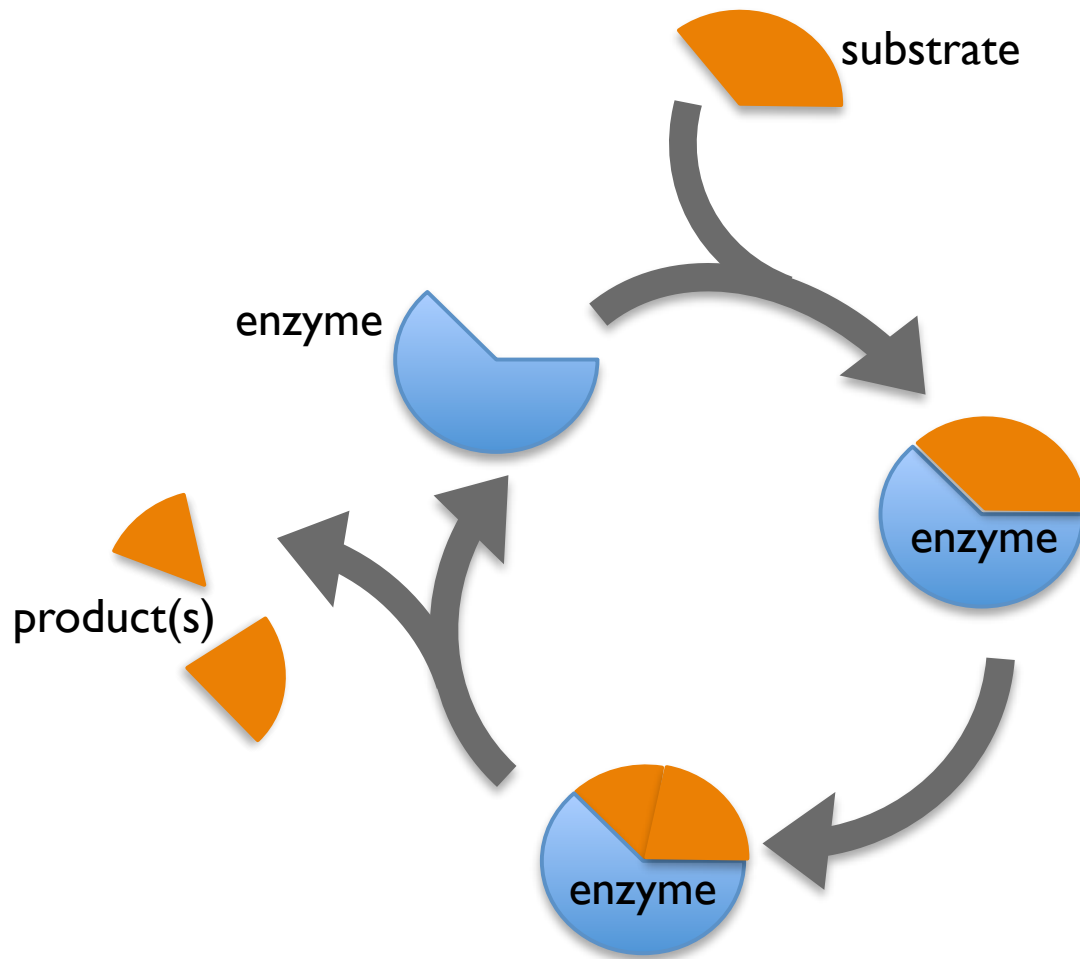
- fluctuations of decay time

$$v_d = \langle T_d \rangle = \langle x_{\max} \rangle / \delta = \alpha\tau / \delta$$

Period variance:

$$v = v_d + v_x / \delta = 2\alpha\tau / \delta$$

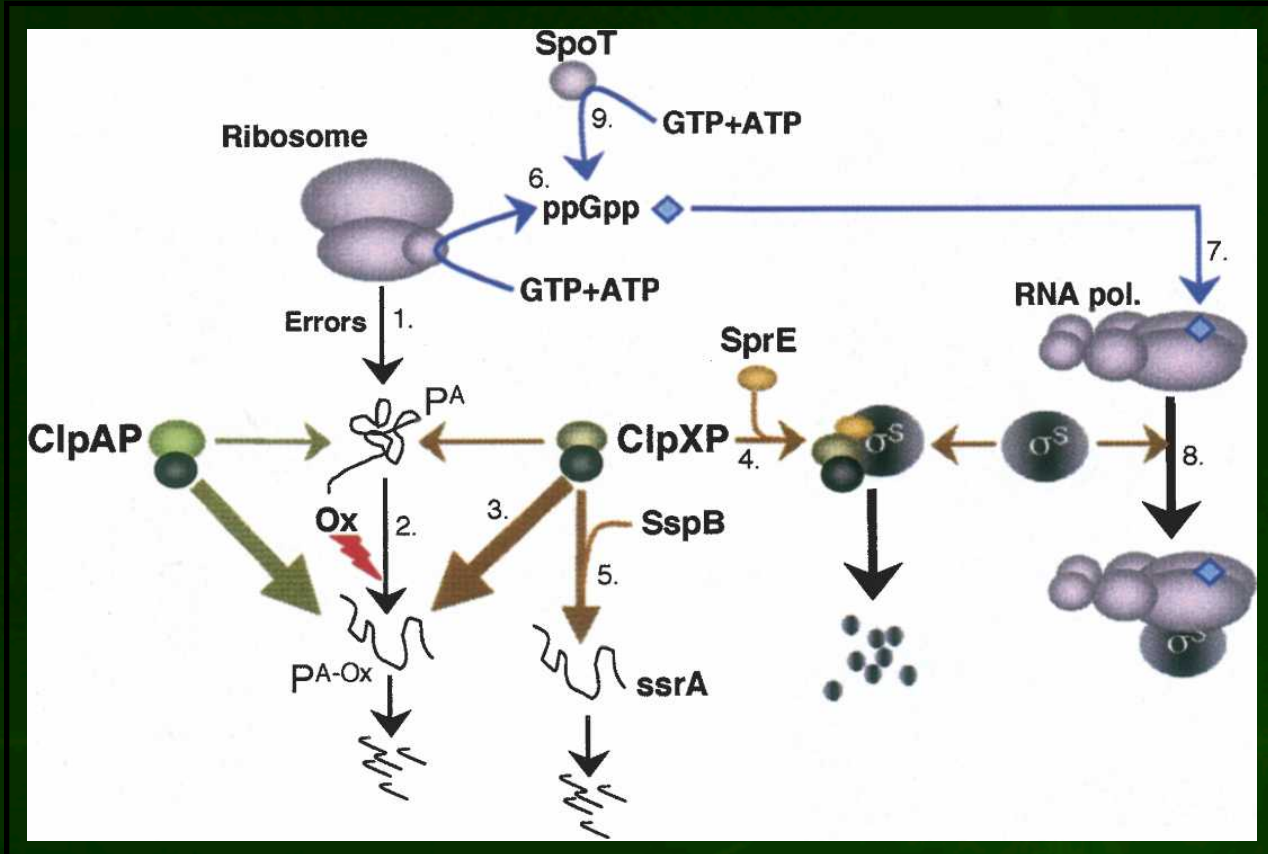
# Enzymatic processing



**Examples:** metabolism, transcription, translation, degradation, signaling...



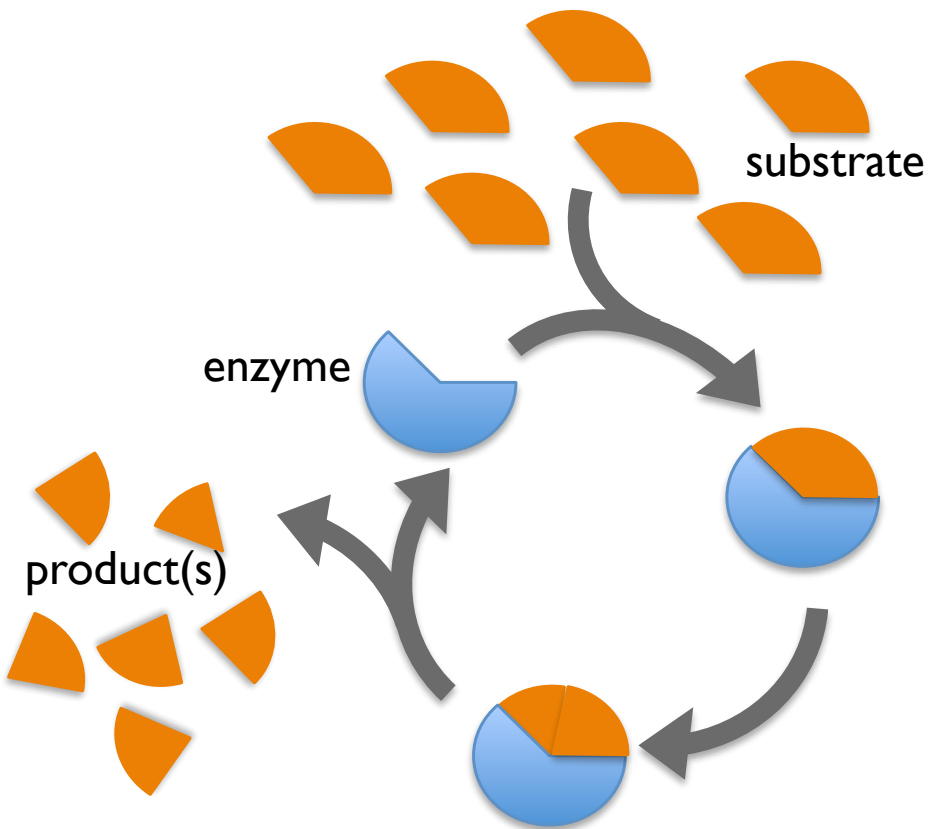
# Enzymatic Queueing: ClpXP degradation machine



ClpXP senses proteins that are tagged for fast degradation and destroys them

normally unstable proteins are stabilized and become active when ClpXP is overloaded

# Enzymatic queueing



Lots of substrate: queues!



How do fluctuations affect queueing?

## Two Cases



Underloaded

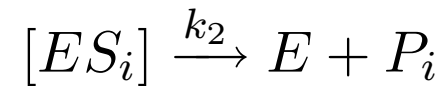
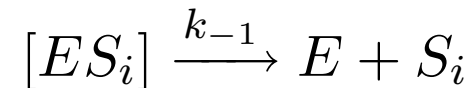
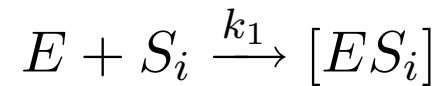
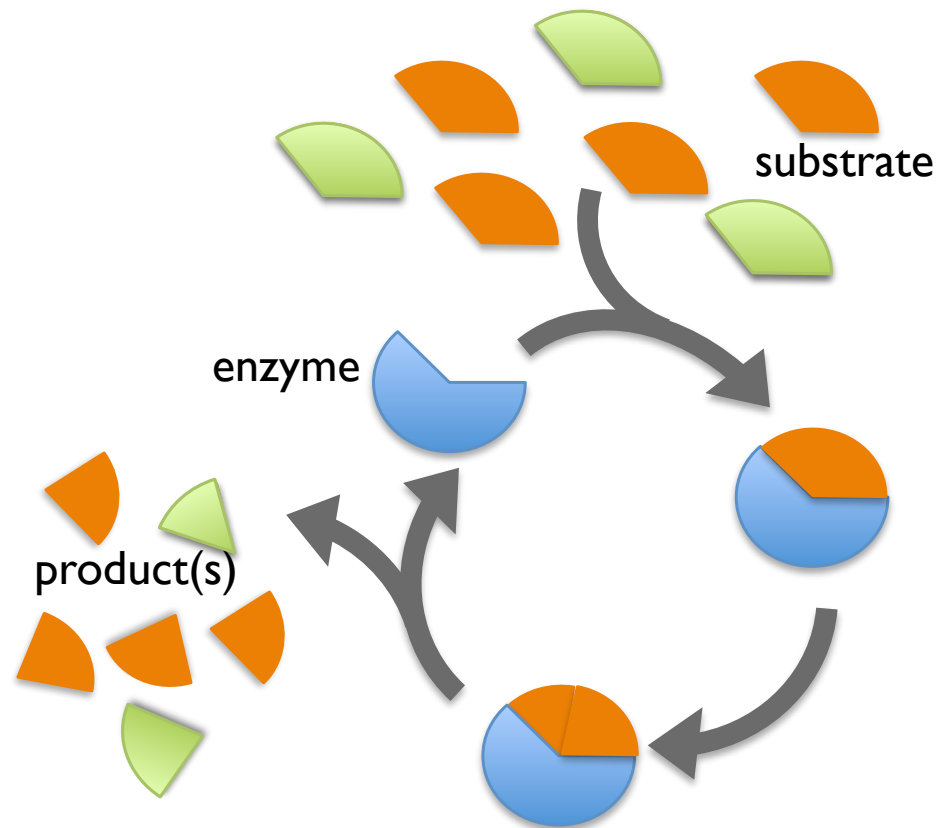
- service rate  $>$  incoming traffic rate
- little competition between customers
- queue lengths are short



Overloaded

- service rate  $<$  incoming traffic rate
- competition between customers
- queue lengths are long

# Enzymatic queueing: multiple substrates



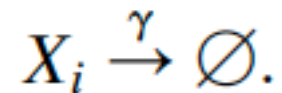
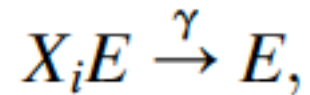
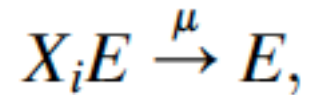
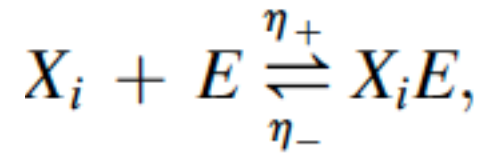
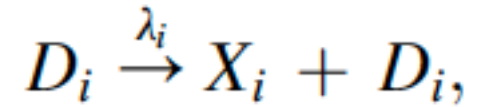
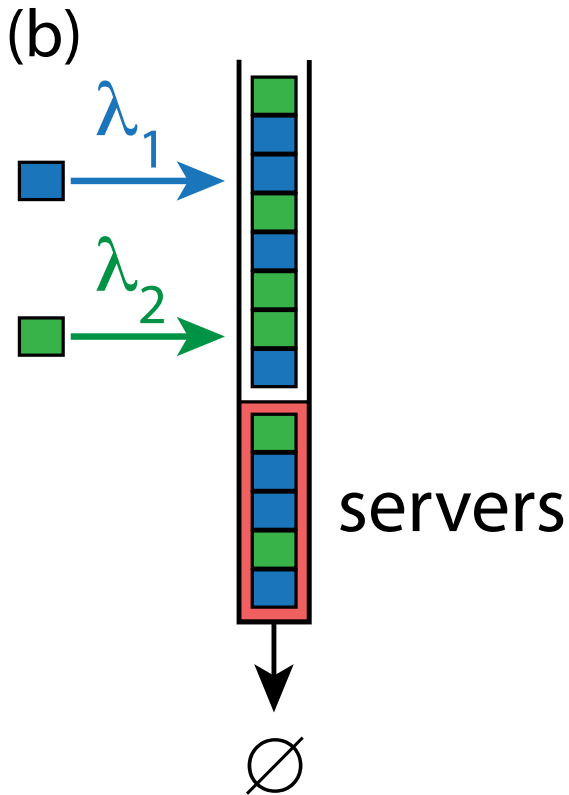
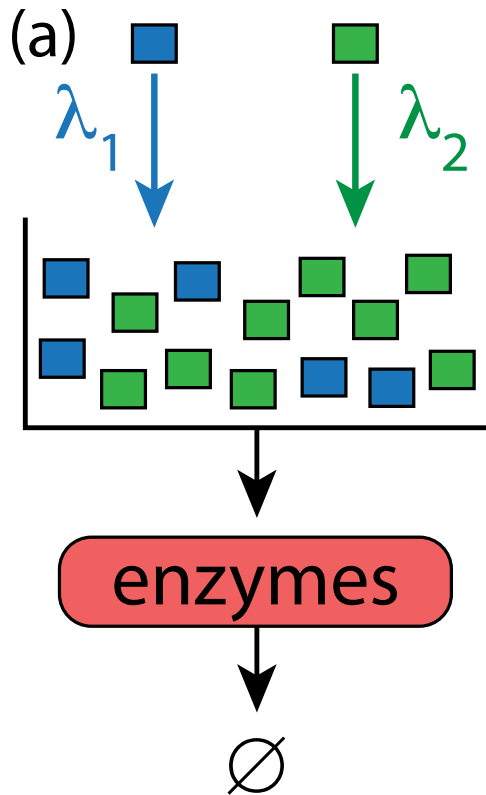
Stochastic kinetics:

fluctuating substrate

fluctuating products

...correlations?

# Queueing Model

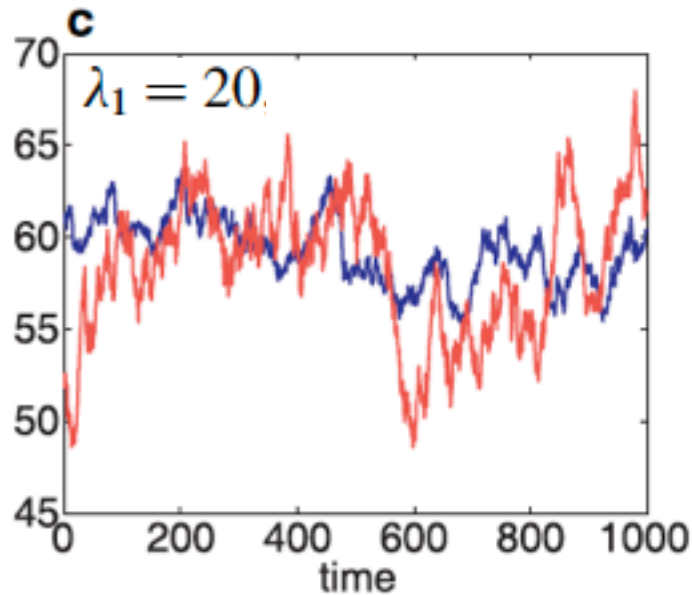
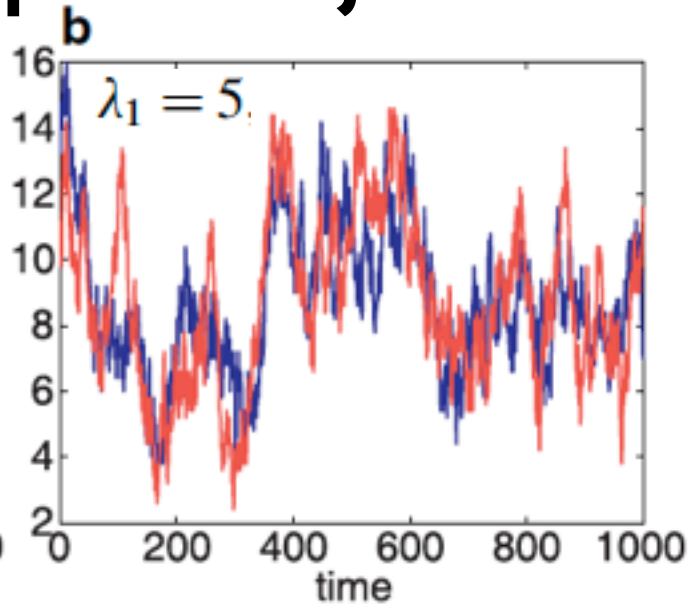
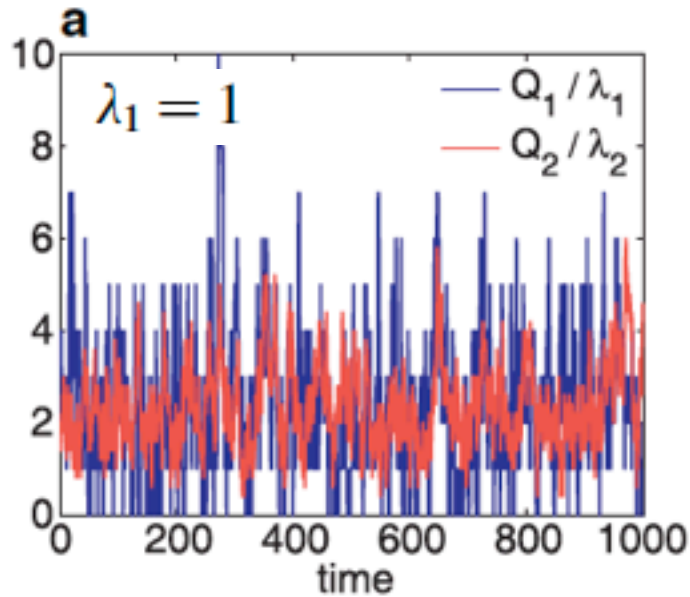


$Q_i(t)$  # of molecules of type  $i$

$$N(t) = \sum_{i=1}^m Q_i(t)$$

$$K = \eta_- / \eta_+$$

# Sample trajectories



$\lambda_2 = 5, \mu = 1, \gamma = 0.01,$   
 $\eta_+ = 20,$  and  $\eta_- = 100$   
number of enzymes  $L = 10$



# Main result

Conditioned on the total number of protein molecules  $N$  in the system being  $n$ , the steady-state distribution for  $Q$  is a multinomial distribution with parameters  $(n; p_1, \dots, p_m)$ , where  $p_i = \lambda_i/\Lambda$ ,  $i = 1, \dots, m$ , and  $\Lambda = \sum_{i=1}^m \lambda_i$

$$P(Q = (q_1, \dots, q_m)) = P(N = n) \frac{n!}{q_1! \dots q_m!} \prod_{i=1}^m p_i^{q_i}$$

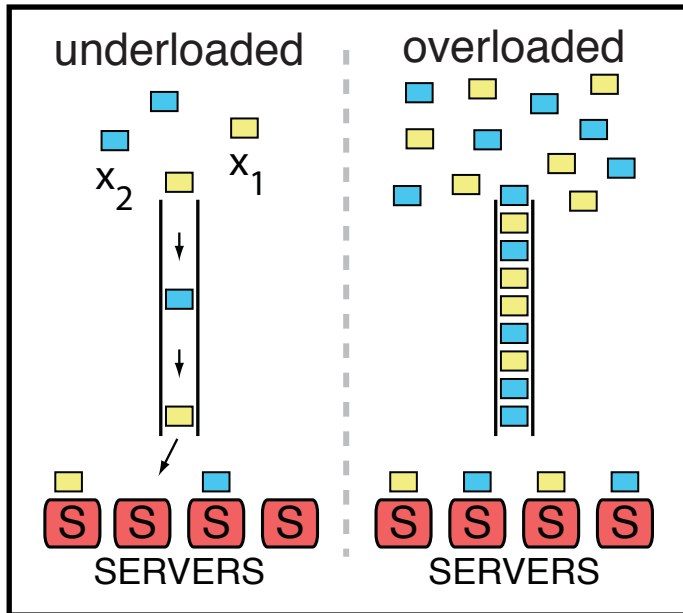
$$P(N = n) = c \frac{\Lambda^n}{\prod_{\ell=1}^n \phi(\ell)}, \quad n = 0, 1, 2, \dots, \quad \phi(n) = \min(n, L)\mu + n\gamma$$

$$n = \sum_{i=1}^m q_i, \quad N = \sum_{i=1}^m Q_i. \quad \nu(N) = \text{Var}(N)/E[N] \text{ is the Fano factor for } N$$

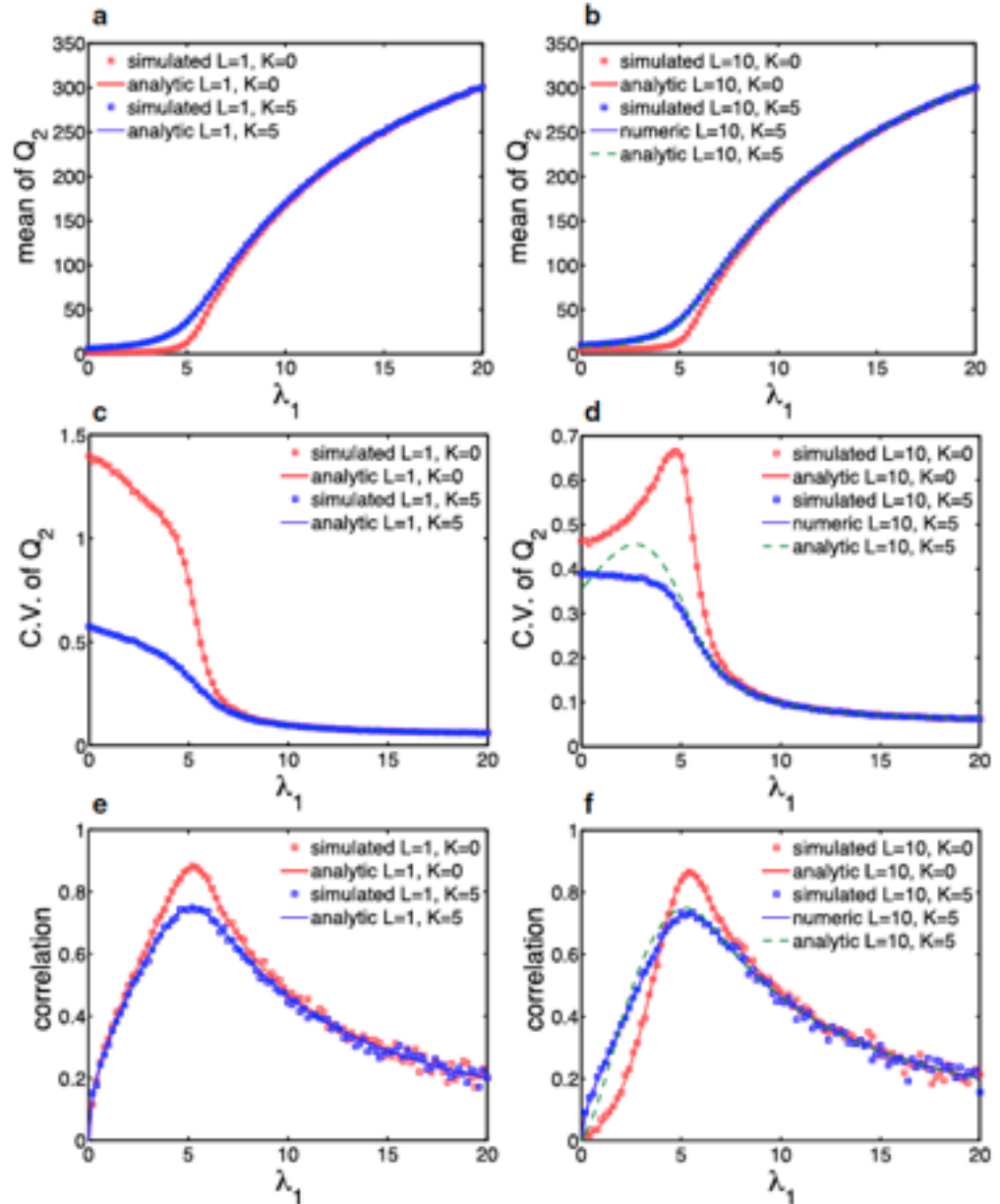
Correlation:

$$r_{ij} = \frac{\nu(N) - 1}{\left(\nu(N) - 1 + \frac{1}{p_i}\right)^{1/2} \left(\nu(N) - 1 + \frac{1}{p_j}\right)^{1/2}}$$

# Queueing Model

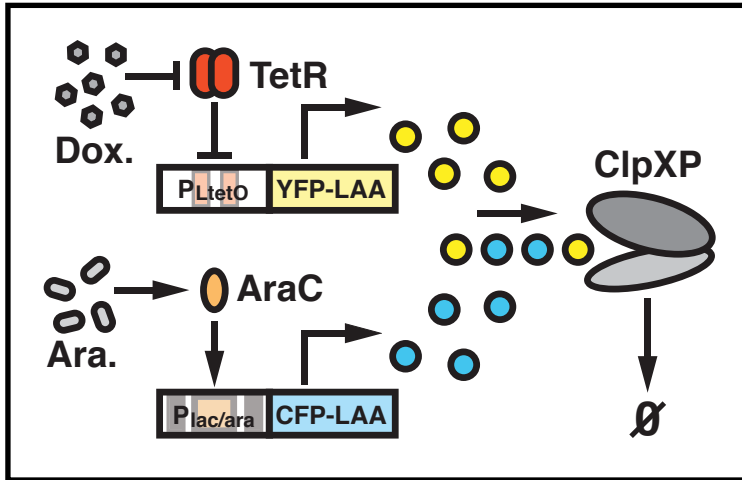


$$K = \eta_- / \eta_+$$



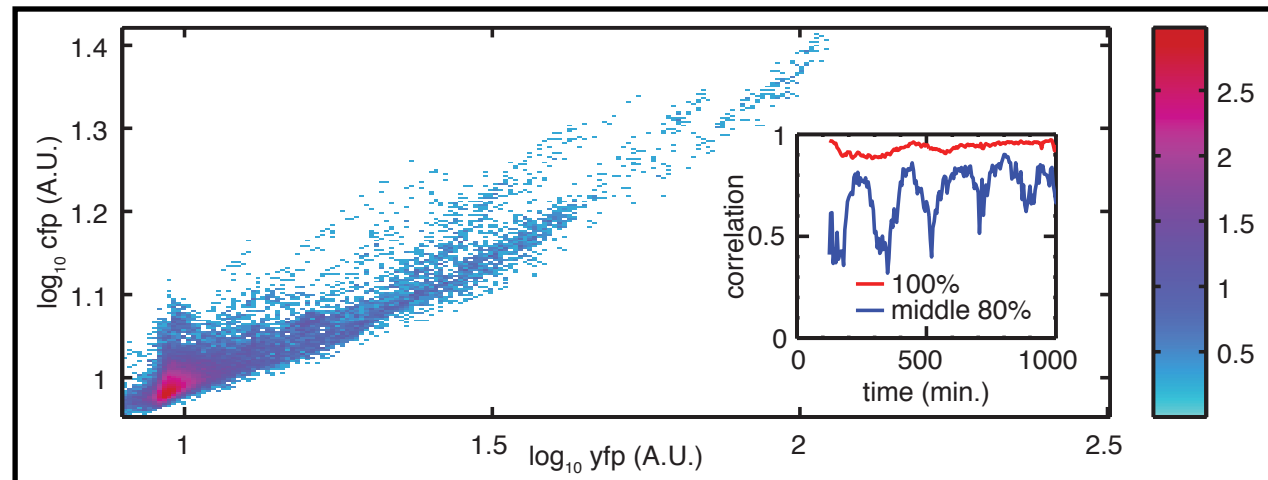
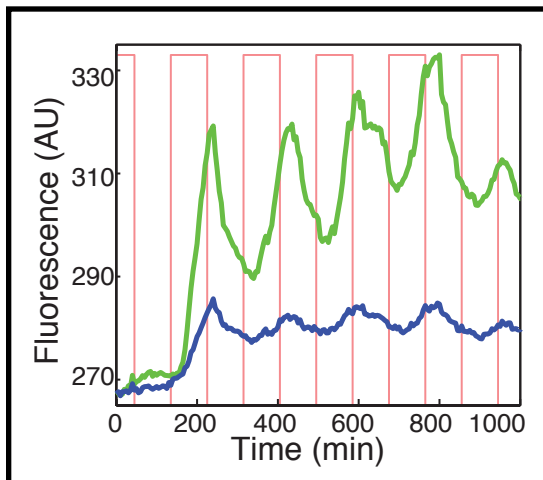


# Queueing in experiment: ClpXP degradation



on a microfluidic device, we drove direct induction (green) by a time-dependent chemical signal

time-dependent signal is carried through



periodic noise resonance associated with crossing through balance

# Conclusions

- Genetic regulation is strongly affected by fluctuations, both intrinsic and extrinsic
  - In many cases, extrinsic factors dominate
  - Theoretical description of extrinsic variability is developed and compared with experimental data from multiple promoter-gene pairs.
  - Deterministic and stochastic description of regulatory delays developed, delays of transcription/translation of auto-repressor may lead to increased fluctuations levels and oscillations even when deterministic model shows no Hopf bifurcation
  - Modified Gillespie algorithm is developed for simulating time-delayed reactions
- 
- D. Volfson, J. Marciniak, N. Ostroff, L. Tsimring, J. Hasty, origins of extrinsic variability in eukaryotic gene expression, *Nature*, 439, 861-864 (16 Feb 2006).
  - D.A. Bratsun, D.N. Volfson, L.S. Tsimring, and J. Hasty, Delay-induced stochastic oscillations in gene regulation. *Proc. Natl. Acad. Sci.*, **102**, no.41, 14593-12598 (2005)
  - J. Stricker, S. Cookson, M. Bennett, W. Mather, L. S. Tsimring, J. Hasty. A robust and tunable synthetic gene oscillator. *Nature*, **456**(7221): 516-9 (2008).
  - W. Mather, M. R. Bennett, J. Hasty, L. S. Tsimring, Delay-induced degrade-and-fire oscillations in small genetic circuits. *Phys. Rev. Lett.*, **102**, 068105 (2009)
  - W. Mather, N. A. Cookson, J. M. Hasty, L. S. Tsimring, and R. J. Williams. Correlation resonance generated by coupled enzymatic processing. *Biophys. J.*, 2010, **99**, 3172-3181