

Effects of extrinsic noise on gene expression and cellular decision making

Michael Assaf 

E. Roberts (Johns Hopkins University)

SIAM "Applications of Dynamical Systems", May 2015 - Snowbird, Utah

Minisymposium: Cellular decision making under noise - SIAM

Outline

- WKB theory of gene expression
- Stochastic gene expression
- Genetic circuits with feedback: genetic switches
- Effects of EN on gene regulation and cellular decision making

WKB theory of stochastic populations

Consider a Poisson process $n \xrightarrow{N} n+1, n \xrightarrow{1} n-1 \Rightarrow \dot{n} = N - n$

$$\dot{P}_n = N(P_{n-1} - P_n) + (n+1)P_{n+1} - nP_n \quad \longrightarrow \quad P_n = \frac{e^{-N} N^n}{n!}$$

WKB theory of stochastic populations

Consider a Poisson process $n \xrightarrow{N} n+1, n \xrightarrow{1} n-1 \Rightarrow \dot{n} = N - n$

$$\dot{P}_n = N(P_{n-1} - P_n) + (n+1)P_{n+1} - nP_n \quad \longrightarrow \quad P_n = \frac{e^{-N} N^n}{n!}$$

↓

$$P_n = e^{-NS(x) - S_1(x) + N^{-1}S_2(x) + \dots}; \quad x \equiv n/N \quad -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + U(x)\psi = E\psi$$

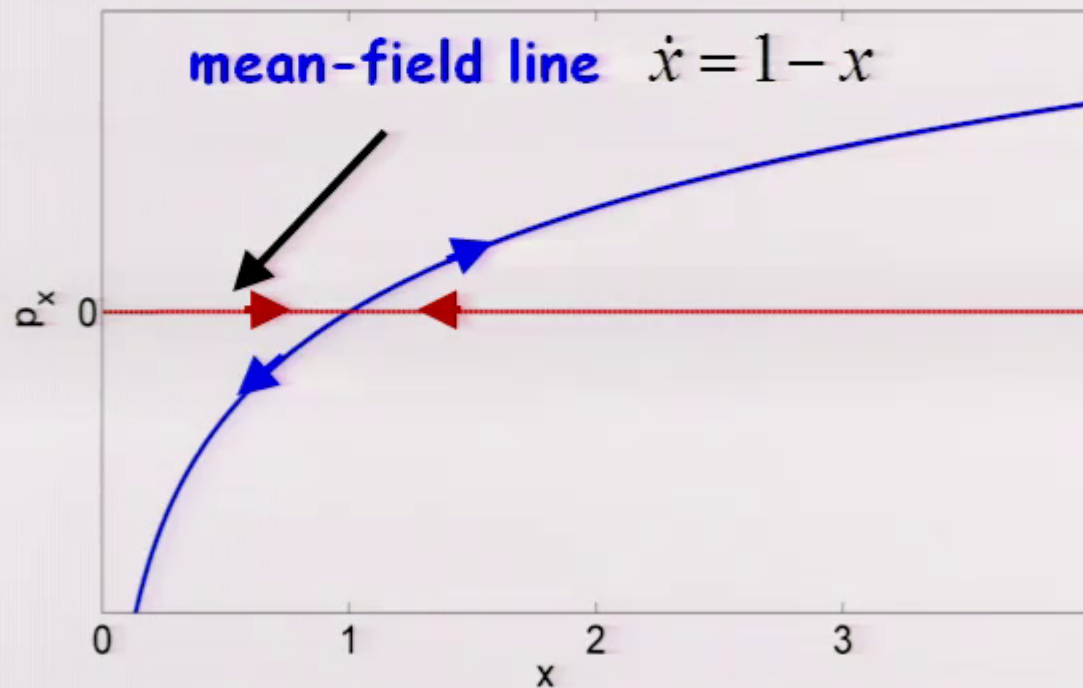
Kubo et. al. 1973, Hu Gang 1987, Dykman et. al. 1994, Kessler and Shnerb 2007, Meerson and Sasorov 2008, Escudero and Kamenev 2009...

Perturbation theory with respect to $N \gg 1$ $N^{-1} \Leftrightarrow \hbar$

leading order $O(N)$: $e^{S'(x)} - 1 + x[e^{-S'(x)} - 1] = 0 \Rightarrow S(x)$

The normalized solution reads: $P(x) \approx (2\pi Nx)^{-1/2} e^{-N(1-x+x \ln x)}$

Leading-order WKB approximation



$$H(x, p_x) = (e^{p_x} - 1)(1 - xe^{-p_x}) = 0$$

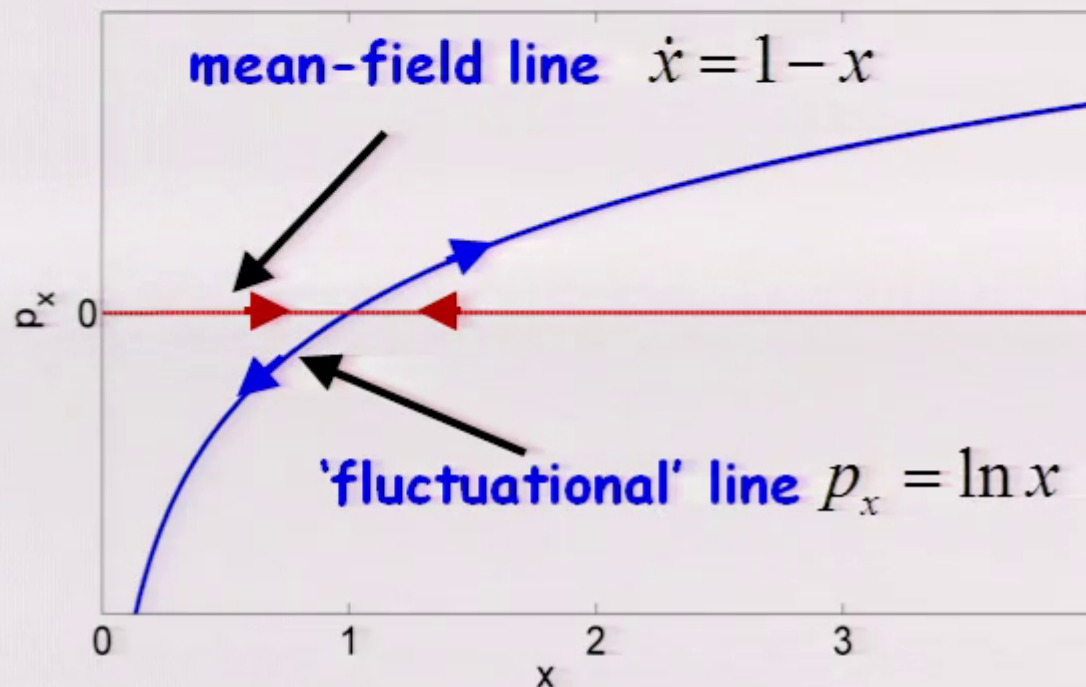
$$p_x(x) \equiv S'(x)$$

'momentum' p_x is a
measure of noise

$$\dot{x} = e^{p_x} - xe^{-p_x}$$

$$\dot{p}_x = 1 - e^{-p_x}$$

Leading-order WKB approximation



$$H(x, p_x) = (e^{p_x} - 1)(1 - xe^{-p_x}) = 0$$

$$p_x(x) \equiv S'(x)$$

'momentum' p_x is a measure of noise

$$\dot{x} = e^{p_x} - xe^{-p_x}$$

$$\dot{p}_x = 1 - e^{-p_x}$$

For $p_x=0$ dynamics goes along mean-field line: $(1,0)$ is an attractor

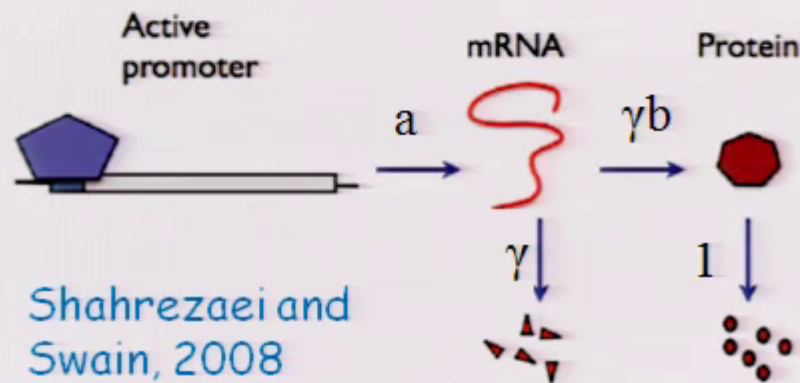
When $p_x \neq 0$, $(1,0)$ becomes saddle and $p_x(x)$ yields stationary PDF

$$S(x) = \int p_x(x) dx = x \ln x - x$$

$$P(x) \sim e^{-NS(x)}$$

Gene expression: unregulated gene

- Proteins are encoded in the DNA in segments that are called genes
- Protein production occurs by transcribing DNA segment into an mRNA via RNA polymerase, which is then translated into a protein by the ribosomes



Rate equations:

$$\dot{M} = a - \gamma M; \quad \dot{N} = \gamma b M - N$$



$$\bar{M} = a/\gamma; \quad \bar{N} = ab$$

a - bursting frequency

b - bursting size

$$\dot{P}_{m,n} = \left[(E_n^1 - 1)n + \gamma(E_m^1 - 1)m + \gamma b m (E_n^{-1} - 1) + a(E_m^{-1} - 1) \right] P_{m,n} ; \quad E_n^j f_k = f_{k+j}$$

Without feedback mRNA is Poisson-distributed

$$P_m = \frac{e^{-a/\gamma}}{m!} \left(\frac{a}{\gamma} \right)^m$$

Alternative description

$$\text{Prob(translation) per mRNA lifetime} = \gamma b / (\gamma b + \gamma) = b / (b + 1)$$

Alternative description

Prob(translation) per mRNA lifetime = $\gamma b / (\gamma b + \gamma) = b / (b + 1)$

proteins synthesized from a given mRNA do not decay before mRNA does \rightarrow they appear simultaneously ($\gamma \gg 1$) Berg (1978)

Prob(r new proteins) = Prob(mRNA translated r times) $P_r = \left(\frac{b}{1+b}\right)^r \left(1 - \frac{b}{1+b}\right)$
 $\langle P_r \rangle = b$

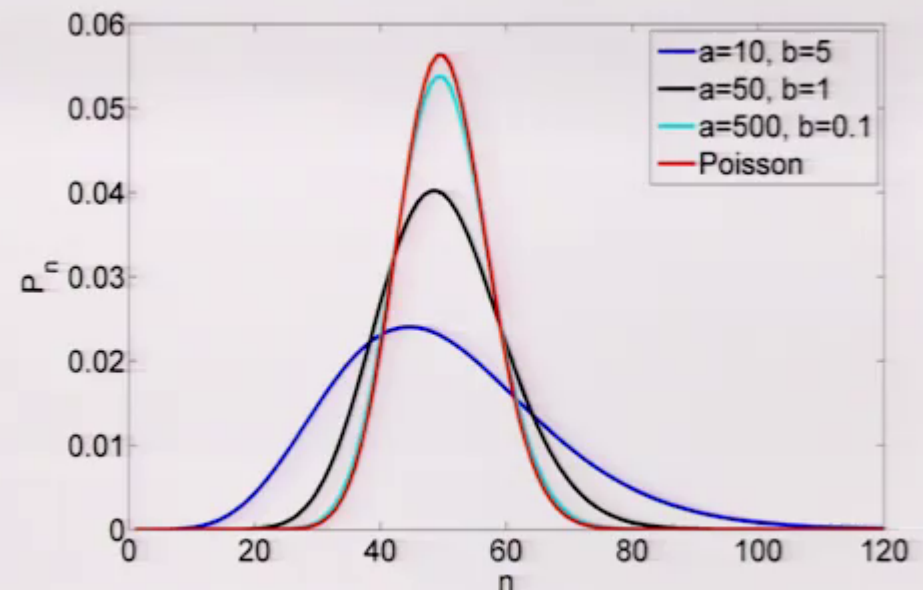
protein production via geometrical bursts

$$\dot{P}_n = a \left[\frac{1}{1+b} \sum_{r=0}^n \left(\frac{b}{1+b}\right)^r P_{n-r} - P_n \right] + (E_n^1 - 1)n$$



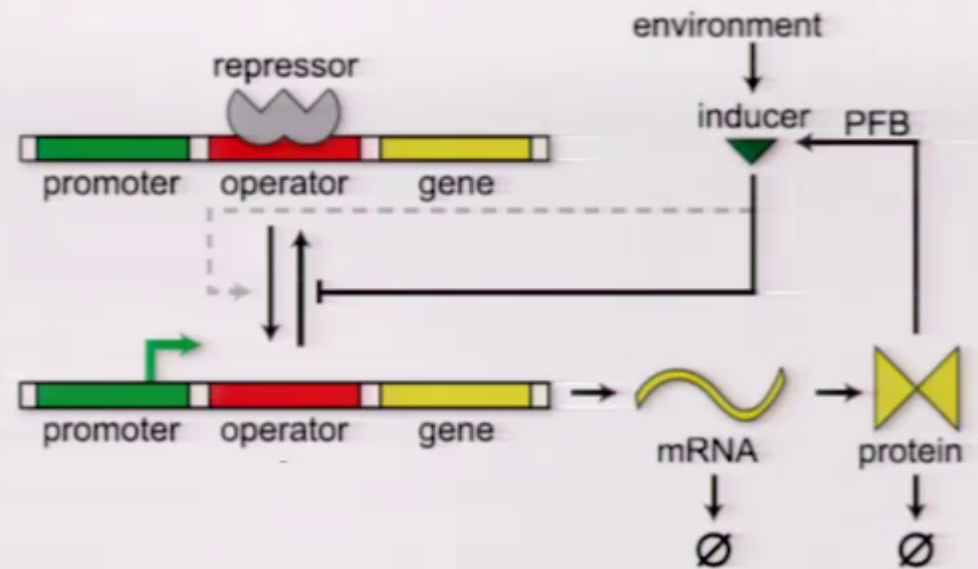
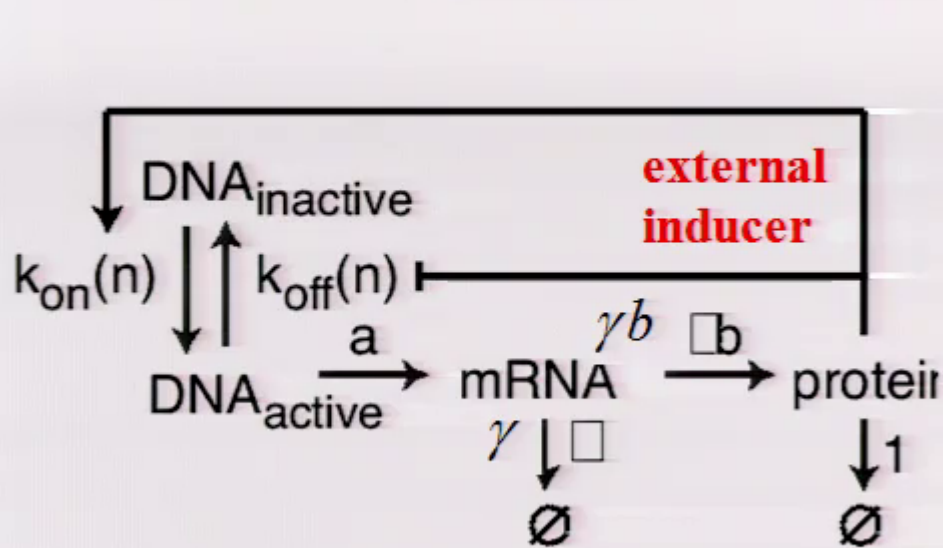
$$P_n = \frac{\Gamma(a+n)}{\Gamma(n+1)\Gamma(a)} \left(\frac{b}{1+b}\right)^n \left(1 - \frac{b}{1+b}\right)^a$$

Paulsson and Ehrenberg (2000)
Friedman, Cai and Xie (2006)



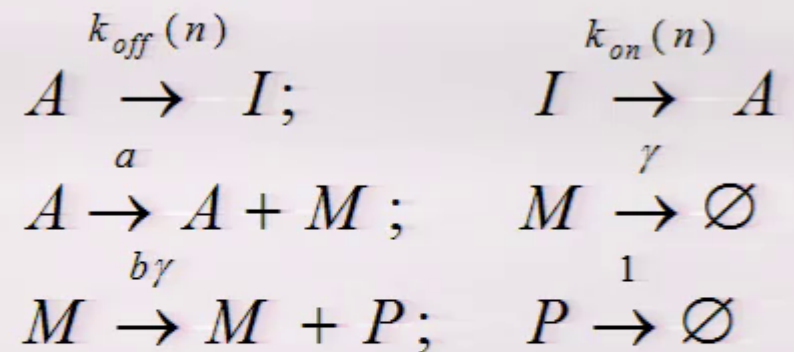
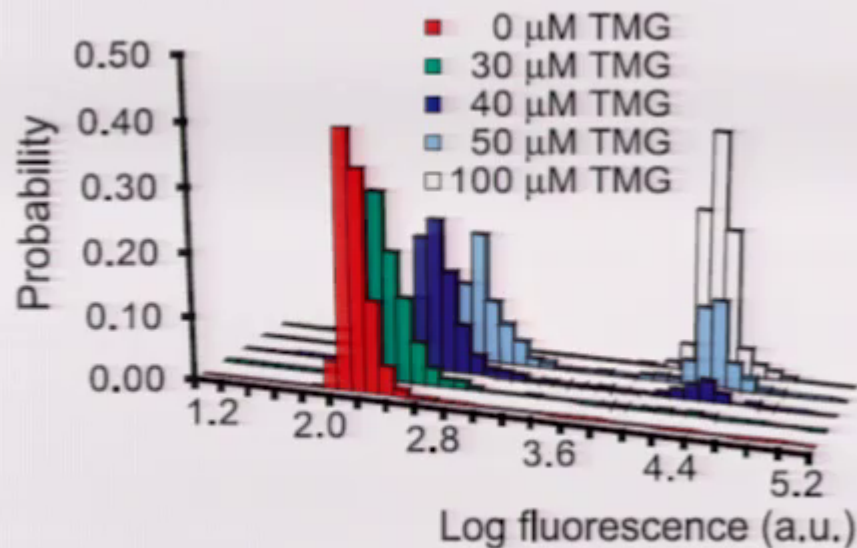
Slide 6 **mRNA fluctuations are important! Variance increases by $(1+b)$**

Gene regulation with feedback: Lac switch



Roberts et al. PLoS Comp. Bio. (2011)

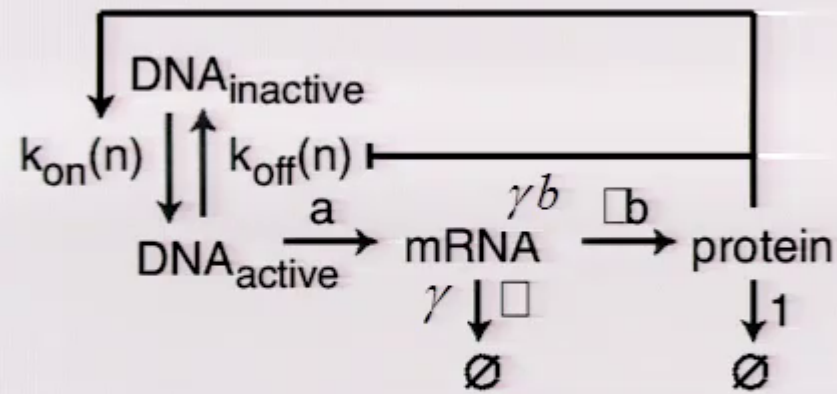
Bimodal distributions



Inducer dynamics fast
controlled by protein

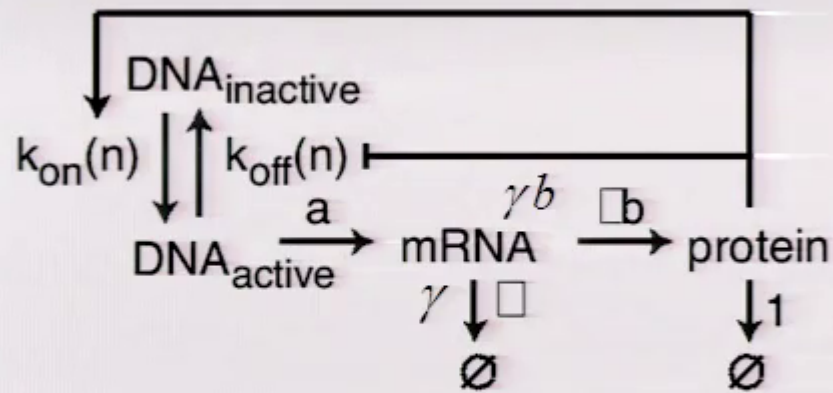
skype
 karkatisha is online

Rate equations

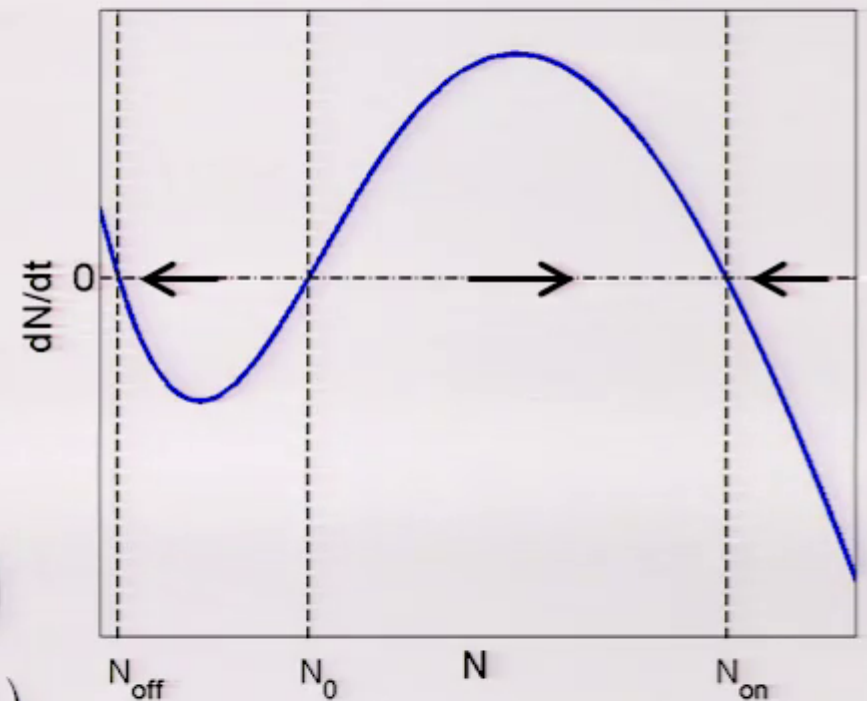
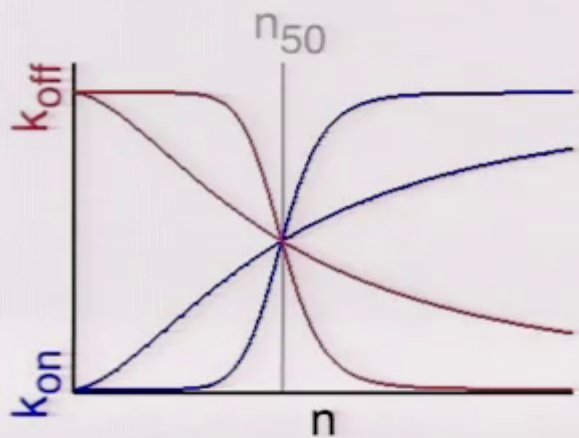


$$\dot{M} = \frac{a k_{\text{on}}(N)}{k_{\text{on}}(N) + k_{\text{off}}(N)} - \gamma M; \quad \dot{N} = \gamma b M - N$$

Rate equations



$$\dot{M} = \frac{a k_{on}(N)}{k_{on}(N) + k_{off}(N)} - \gamma M; \quad \dot{N} = \gamma b M - N$$

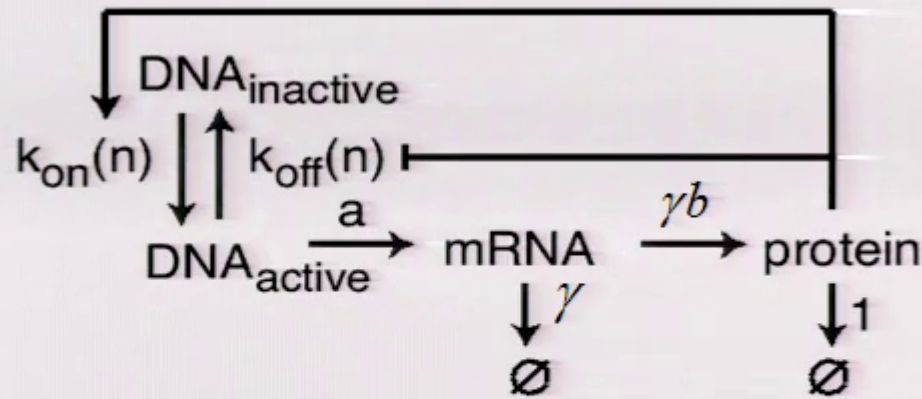


$$k_{on}(n) = k_{min}^0 + (k_{max}^0 - k_{min}^0) n^{h_1} / (n^{h_1} + n_{50}^{h_1})$$

$$k_{off}(n) = k_{max}^1 - (k_{max}^1 - k_{min}^1) n^{h_2} / (n^{h_2} + n_{50}^{h_2})$$

$k_{on}(n)$ and $k_{off}(n)$ are Hill-functions of inducer concentration

Master equation



$P_{m,n} / Q_{m,n}$ probability to find m mRNAs and n proteins for active/inactive DNA

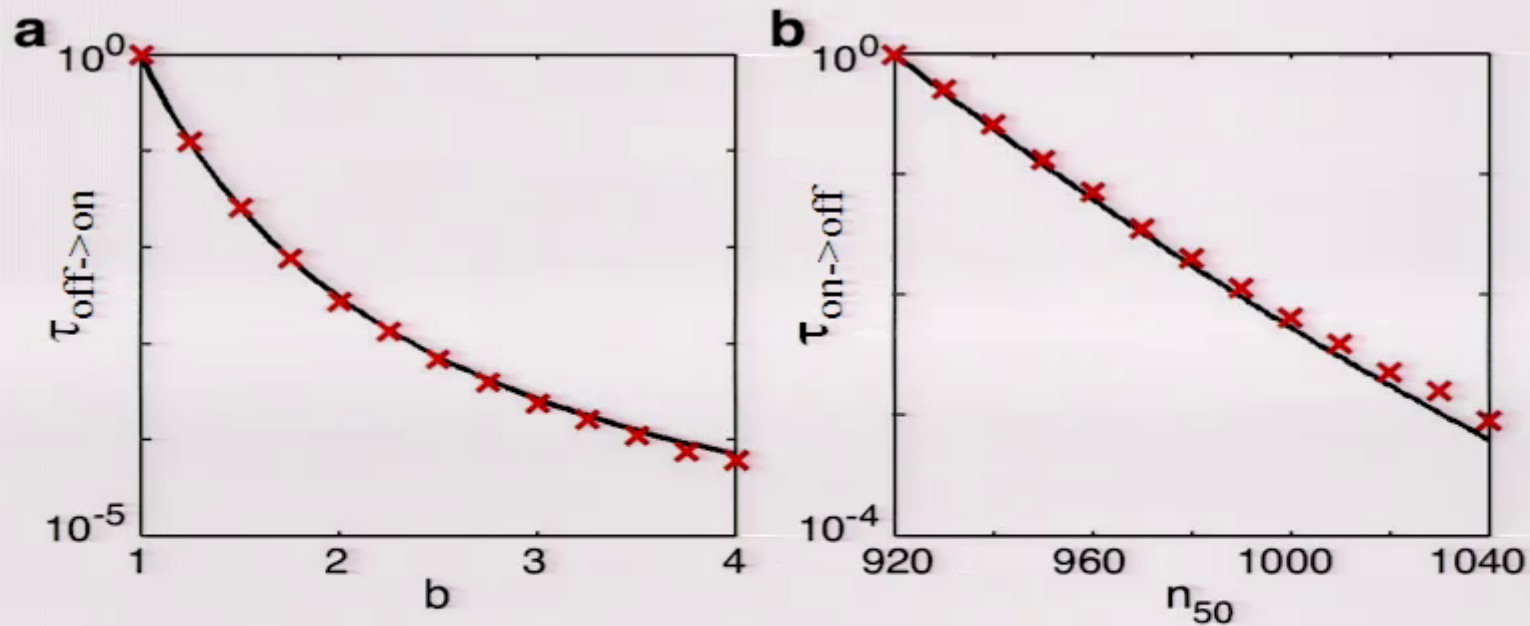
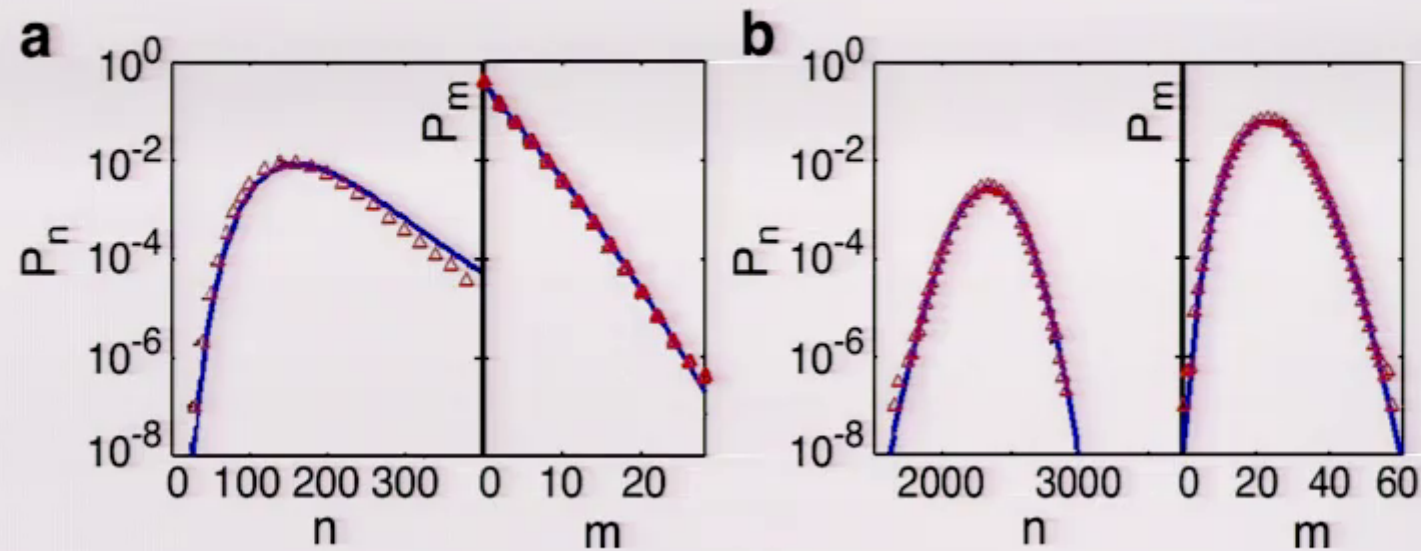
$$\dot{Q}_{m,n} = -K_{off}(n)Q_{m,n} + K_{on}(n)P_{m,n} + \left[a(E_m^{-1} - 1) + \gamma(E_m^1 - 1)m + \gamma b m(E_n^{-1} - 1) + (E_n^1 - 1)n \right] Q_{m,n}$$

$$\dot{P}_{m,n} = K_{off}(n)Q_{m,n} - K_{on}(n)P_{m,n} + \left[\gamma(E_m^1 - 1)m + \gamma b m(E_n^{-1} - 1) + (E_n^1 - 1)n \right] P_{m,n}$$

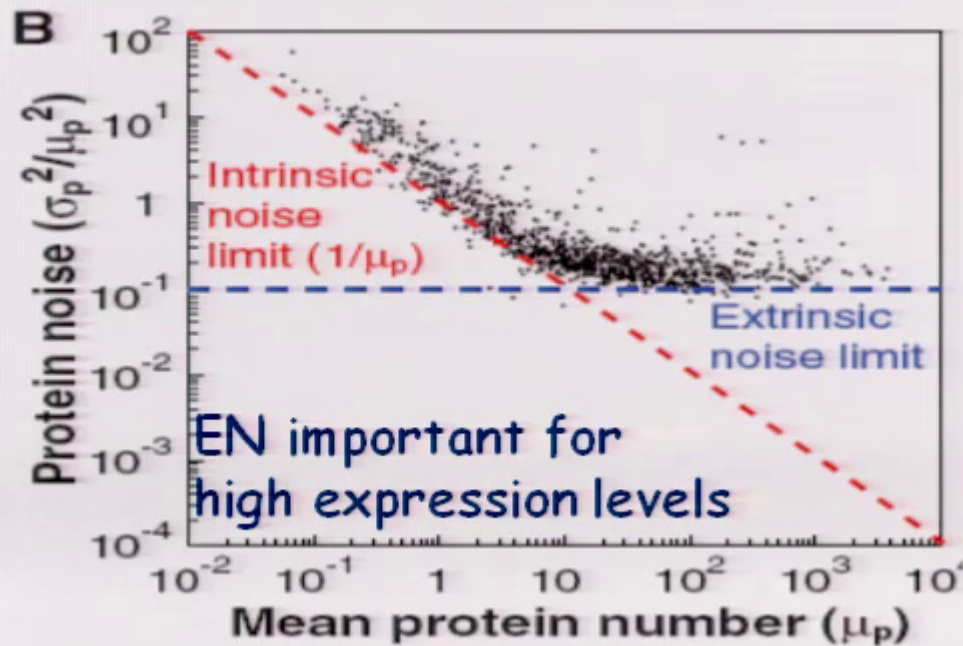
We use WKB theory $P_{m,n} \sim e^{-NS(x,y)}$; $x = m/N$, $y = n/N$, $N = ab$

$\gamma \gg 1$ **adiabatic elimination of mRNA species**

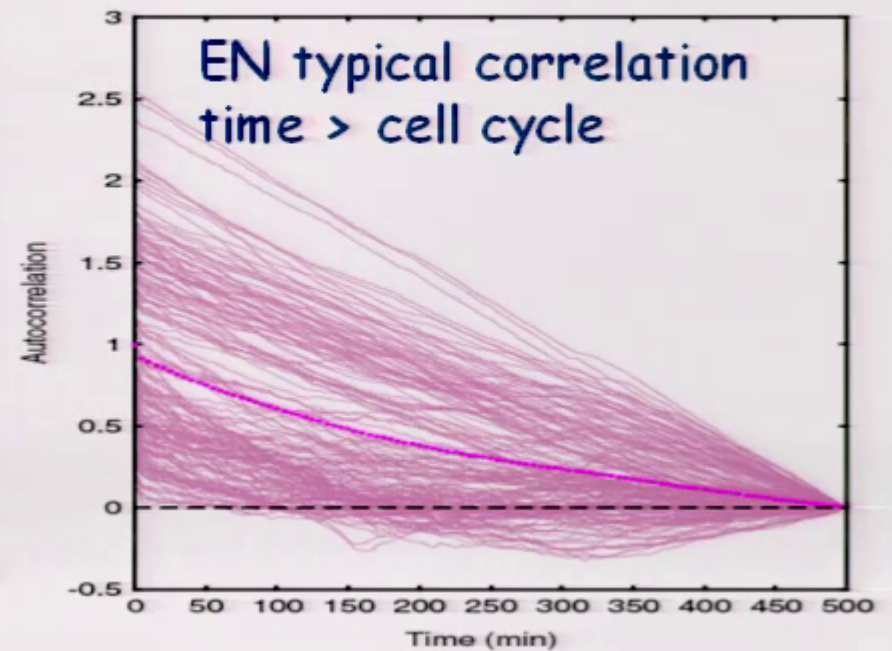
mRNA and protein PDFs and mean switching times



This treatment ignores extrinsic noise and cell to cell variability

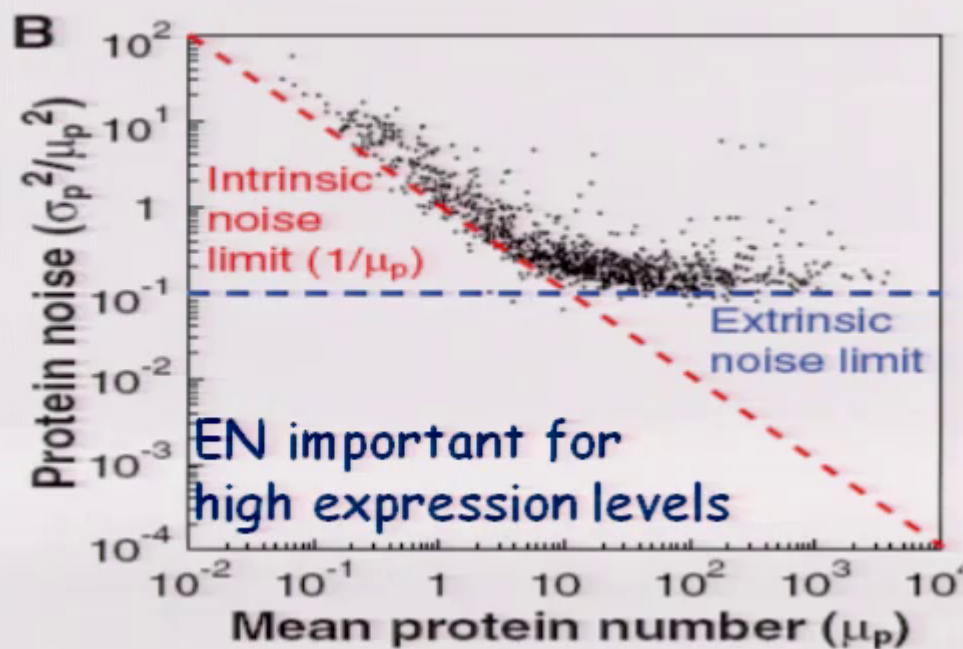


(B) Protein expression noise (σ_p^2/μ_p^2) versus the mean copy number per cell (μ_p). When $\mu_p < 10$, protein expression noise is close to the intrinsic noise limit, which is inversely proportional to the mean (red dashed line). When $\mu_p > 10$, noise becomes independent of the mean and is above a plateau of ~ 0.1 (blue dashed lines), which is the extrinsic noise limit.

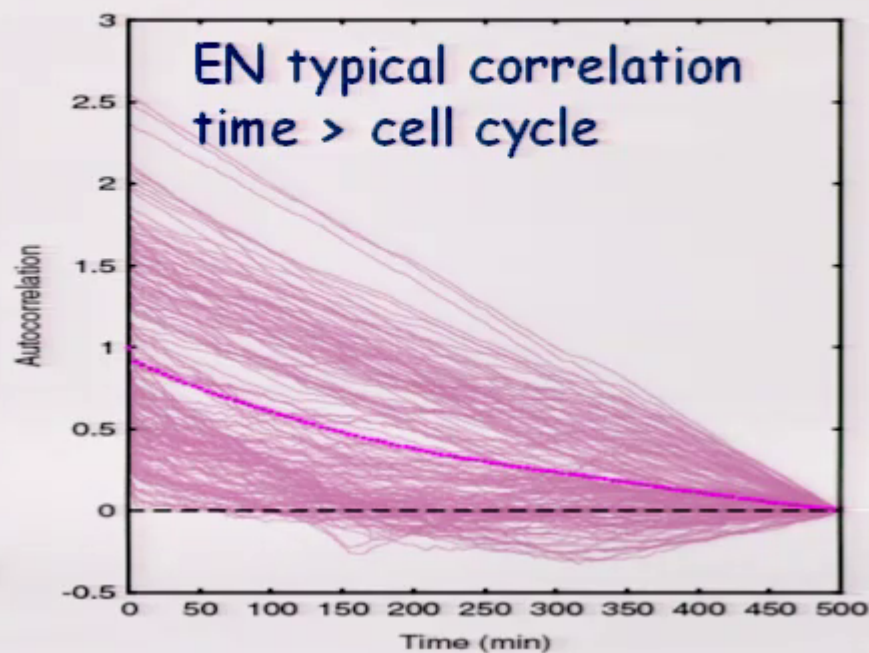


Taniguchi et. al., Science (2010)

This treatment ignores extrinsic noise and cell to cell variability



(B) Protein expression noise (σ_p^2/μ_p^2) versus the mean copy number per cell (μ_p). When $\mu_p < 10$, protein expression noise is close to the intrinsic noise limit, which is inversely proportional to the mean (red dashed line). When $\mu_p > 10$, noise becomes independent of the mean and is above a plateau of ~ 0.1 (blue dashed lines), which is the extrinsic noise limit.



Taniguchi et. al., Science (2010)

Reasons why extrinsic noise (EN) has to be in model:

- fluctuations in environment
- cells are not identical (num of ribosomes, doubling time)
- System is coupled to other fluctuating entities

Adding extrinsic noise

We add EN to one of the reaction rates

$$\alpha \rightarrow \alpha(t) = \alpha_0 + \xi(t)$$

$\xi(t)$ - extrinsic noise variable

$$\langle \xi(t)\xi(t') \rangle = \sigma^2 e^{-|t-t'|/\tau_c}$$

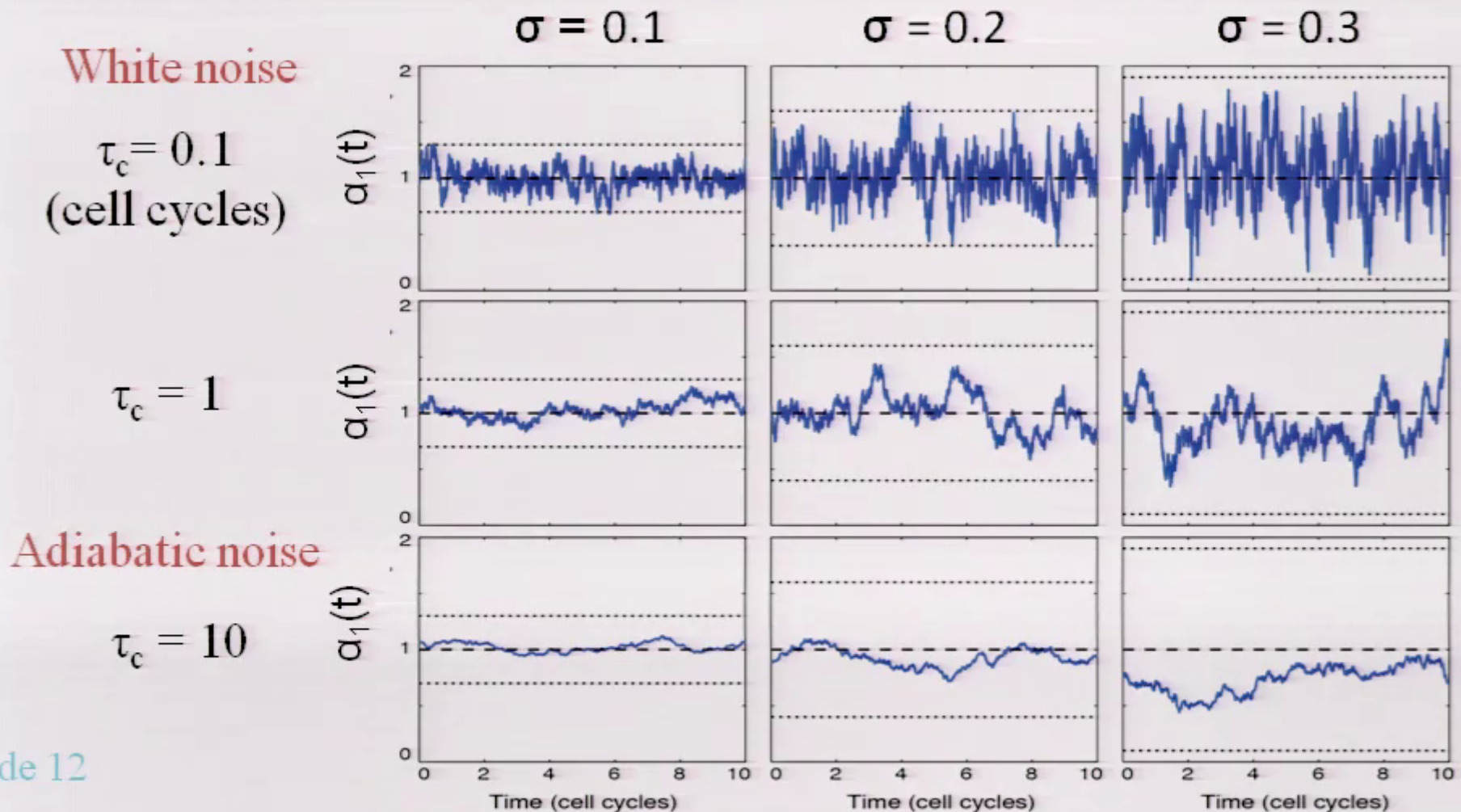
Adding extrinsic noise

We add EN to one of the reaction rates

$$\alpha \rightarrow \alpha(t) = \alpha_0 + \xi(t)$$

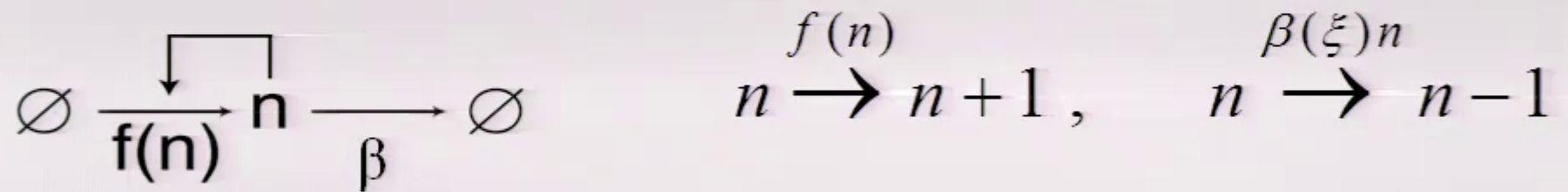
$\xi(t)$ - extrinsic noise variable

$$\langle \xi(t)\xi(t') \rangle = \sigma^2 e^{-|t-t'|/\tau_c}$$



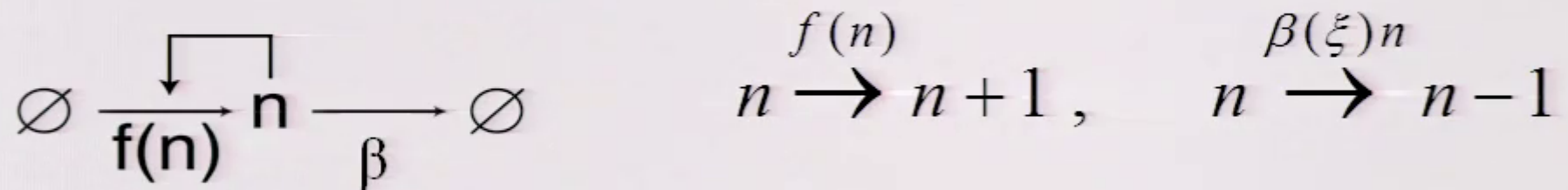
Adding extrinsic noise

Reaction rates depend on auxiliary extrinsic noise variable ξ



Adding extrinsic noise

Reaction rates depend on auxiliary extrinsic noise variable ξ



$P(\xi)$ - statistics of extrinsic noise

$$\beta \rightarrow 1 + \xi(t)$$

$$\langle \xi \rangle \equiv \mu$$

$$\text{var}(\xi) = \sigma^2$$

$$\langle \xi(t)\xi(t') \rangle = \sigma^2 e^{-|t-t'|/\tau_c}$$



$$\dot{P}_{n,\xi} = f(n-1)P_{n-1} - f(n)P_n + (1+\xi)[(n+1)P_{n+1} - nP_n] + \text{dynamics of } \xi$$

How does protein PDF width depend on EN?

Substituting $P_{n,\xi} = P(x,\xi) \sim e^{-NS(x,\xi)}$

for intrinsic noise only

$$\text{var}(n) = \frac{Nx_*}{1 - f'(x_*)}$$

x_* - stable fixed point

$$V = N\sigma^2$$

How does protein PDF width depend on EN?

Substituting $P_{n,\xi} = P(x,\xi) \sim e^{-NS(x,\xi)}$

In the white-noise limit $\tau_c \ll 1$

$$\text{var} = \frac{Nx_*}{1 - f'(x_*)} (1 + x_* V \tau_c)$$

for intrinsic noise only

$$\text{var}(n) = \frac{Nx_*}{1 - f'(x_*)}$$

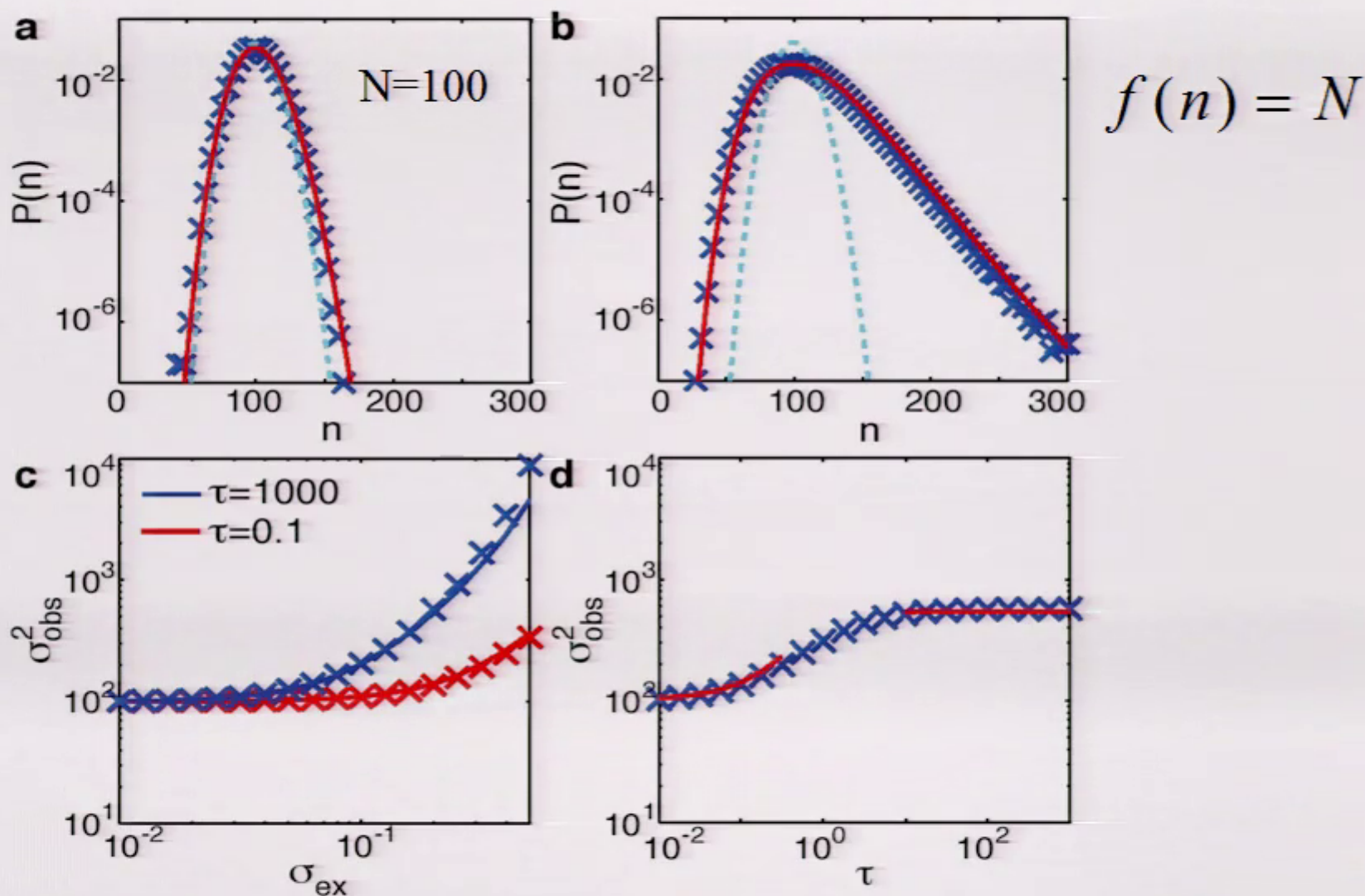
In the adiabatic limit $\tau_c \gg 1$

x_* - stable fixed point

$$\text{var} = \frac{Nx_*}{1 - f'(x_*)} \left[1 + \frac{Vx_*}{1 - f'(x_*)} \right]$$

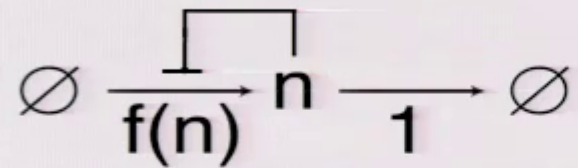
$$V = N\sigma^2$$

Example: non-regulated gene



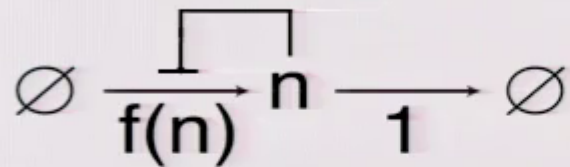
PDF width increased by $1+V\tau_c$ for white EN and $1+V$ for adiabatic EN

Self-inhibitory gene



$$\dot{n} = \frac{N(1+\beta)}{1+\beta(n/N)^h} - \alpha(t)n$$

Self-inhibitory gene



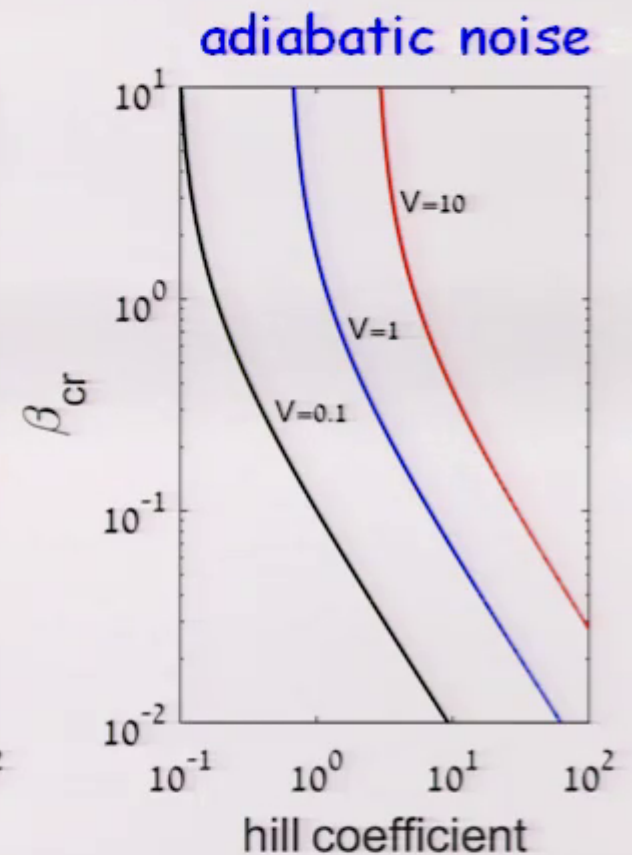
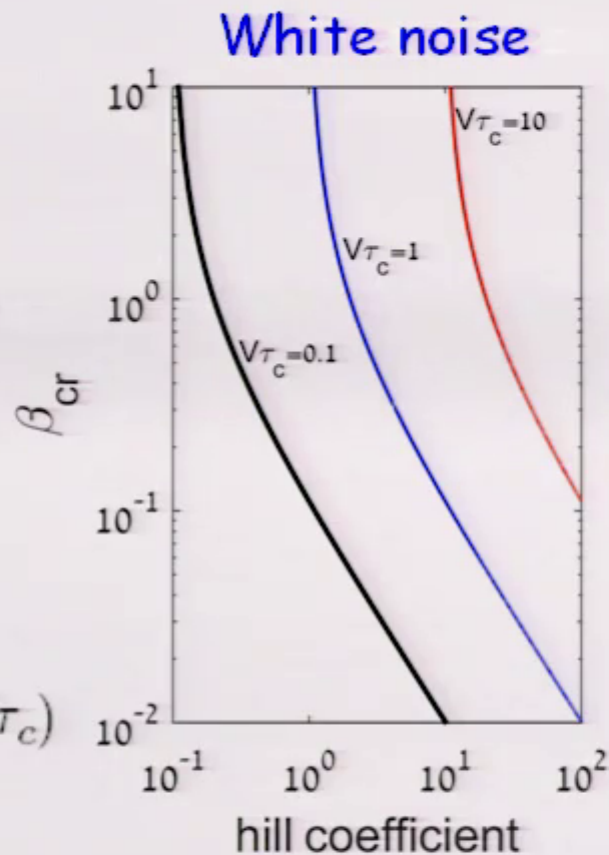
$$\dot{n} = \frac{N(1+\beta)}{1+\beta(n/N)^h} - \alpha(t)n$$

white noise

$$\text{var} = \frac{N(1+\beta)}{1+\beta(h+1)} (1+V\tau_c)$$

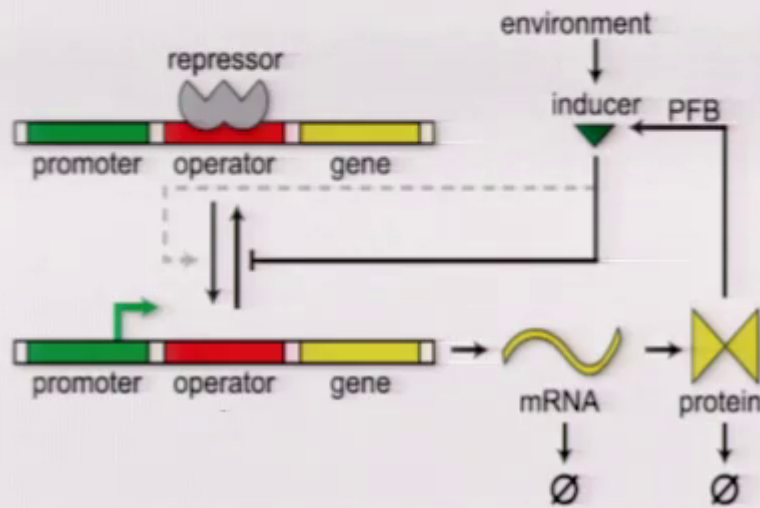
adiabatic noise

$$\text{var} = \frac{N(1+\beta)}{1+\beta(h+1)} \left[1 + \frac{V(1+\beta)}{1+\beta(h+1)} \right]$$

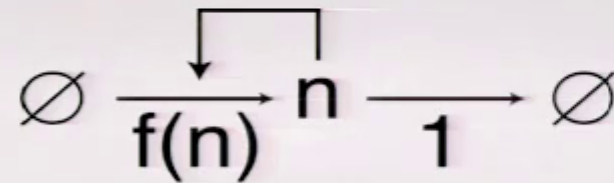


Inhibition can filter EN if accurate cellular decisions are required

How do switching times depend on EN?

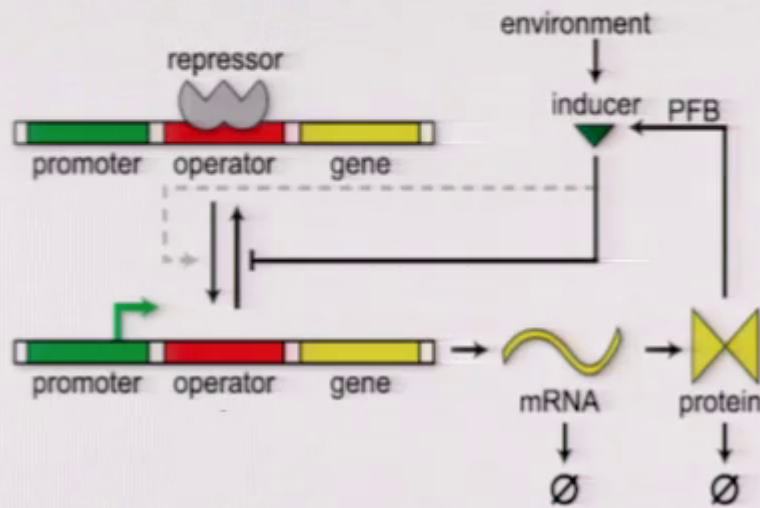


Self-regulating gene (SRG) model

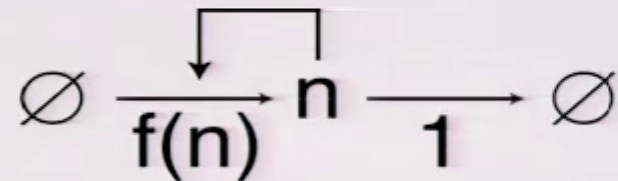


$$\dot{x} = f(x) - x; \quad f(x) = \alpha_0 + \frac{(1 - \alpha_0)x^h}{x^h + x_0^h}$$

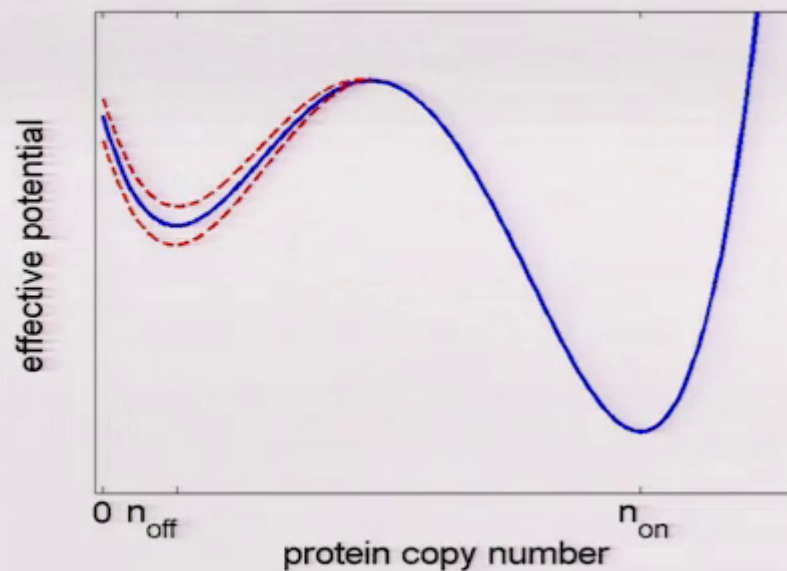
How do switching times depend on EN?



Self-regulating gene (SRG) model

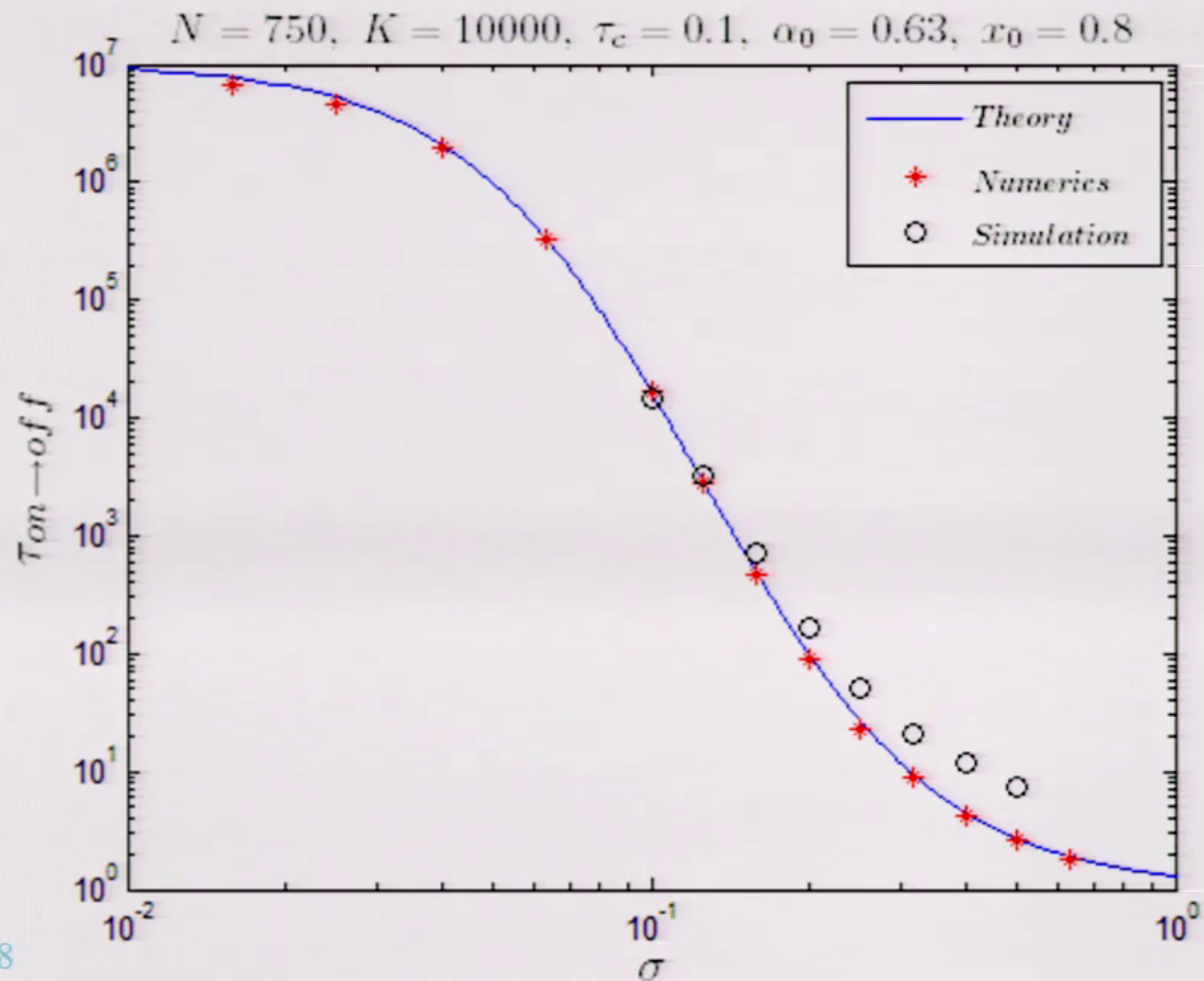


$$\dot{x} = f(x) - x; \quad f(x) = \alpha_0 + \frac{(1 - \alpha_0)x^h}{x^h + x_0^h}$$

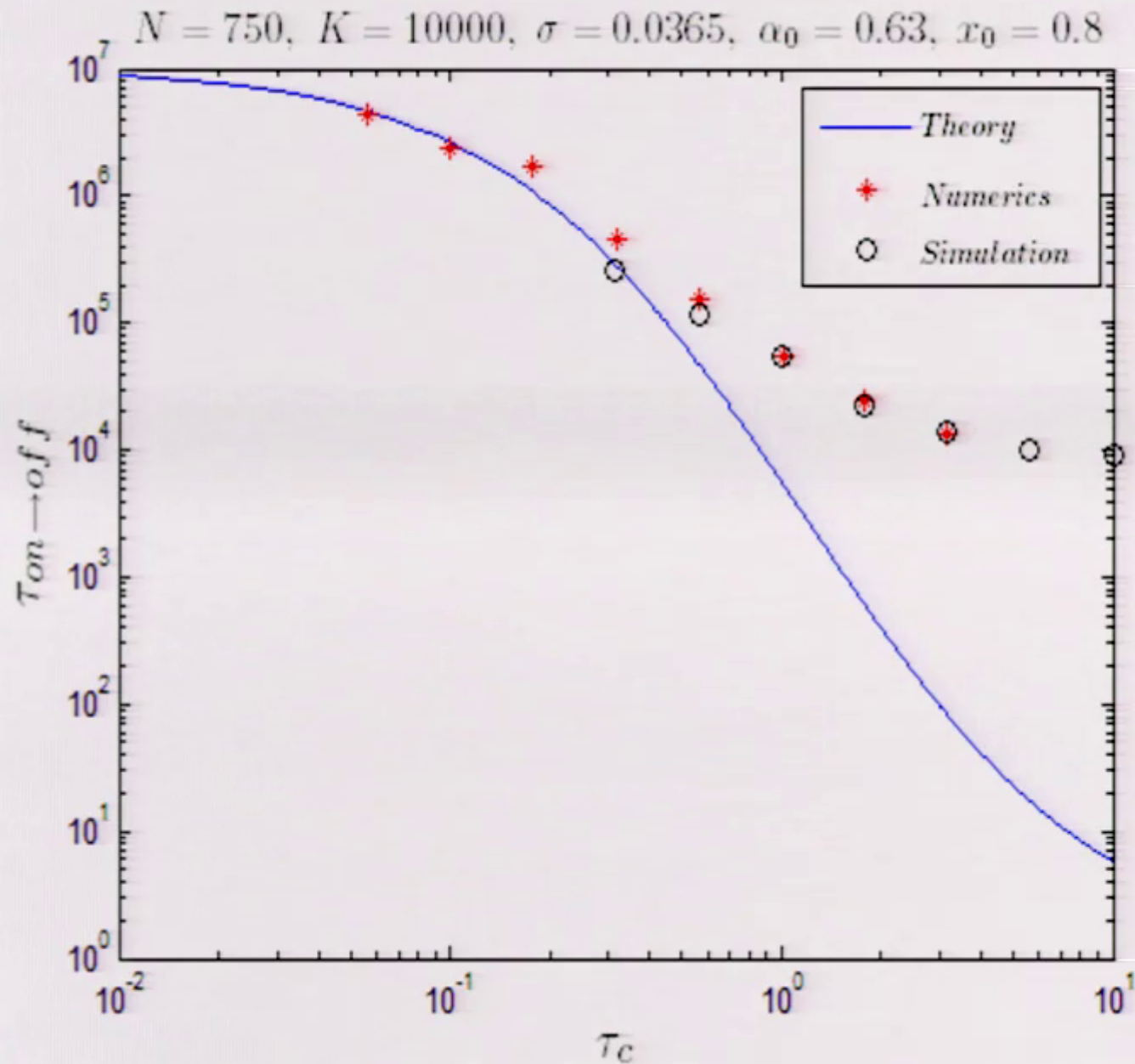


System selects "easiest" path to switch

EN exponentially decreases switching times



Dependence of switching time on τ_c is non-monotonic

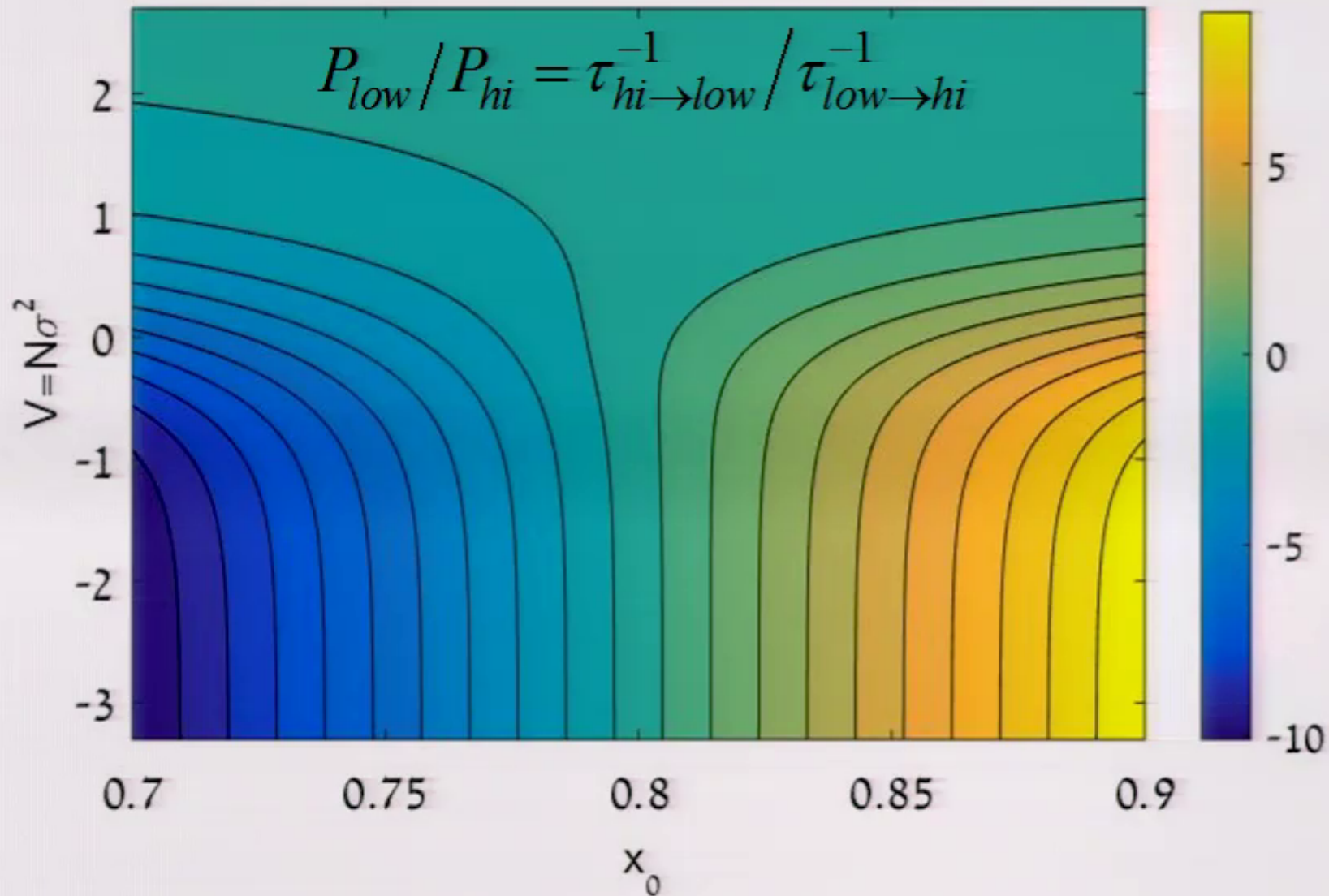


Evidence of optimal τ_c that minimizes switching time and maximizes heterogeneity (may be achieved by tuning division times to match τ_c)

At strong EN switching is EN-driven

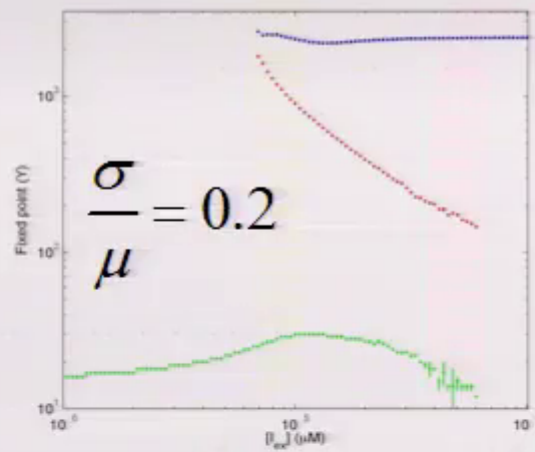
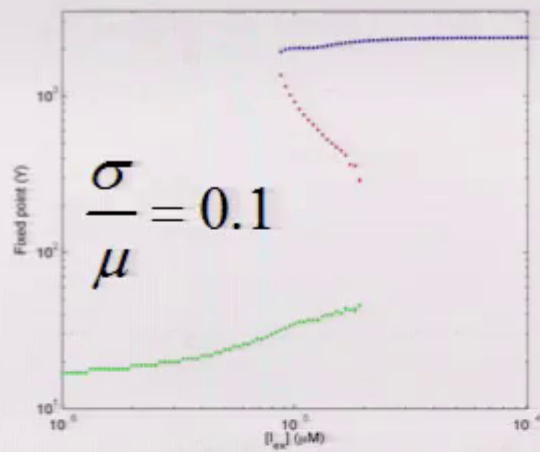
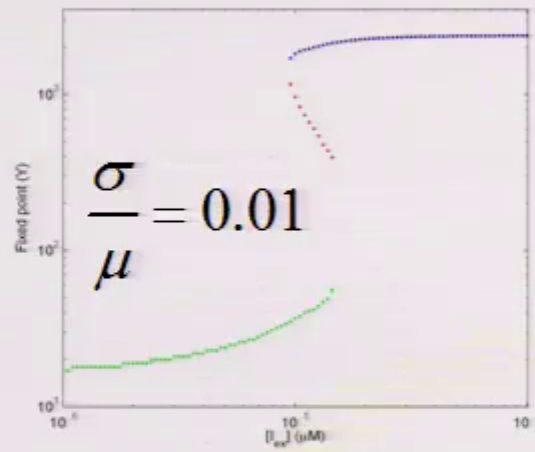
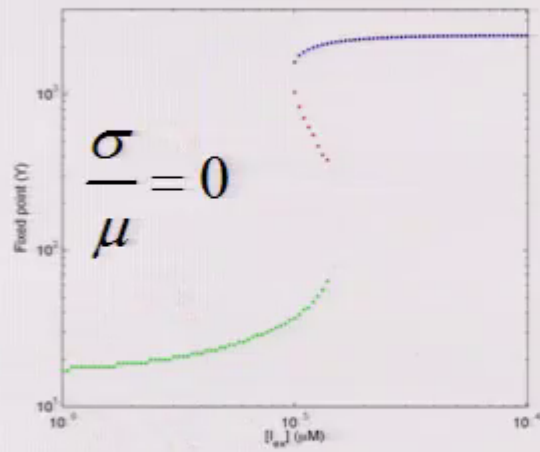
$$\tau_{switching} \sim \tau_c$$

EN significantly alters switching times

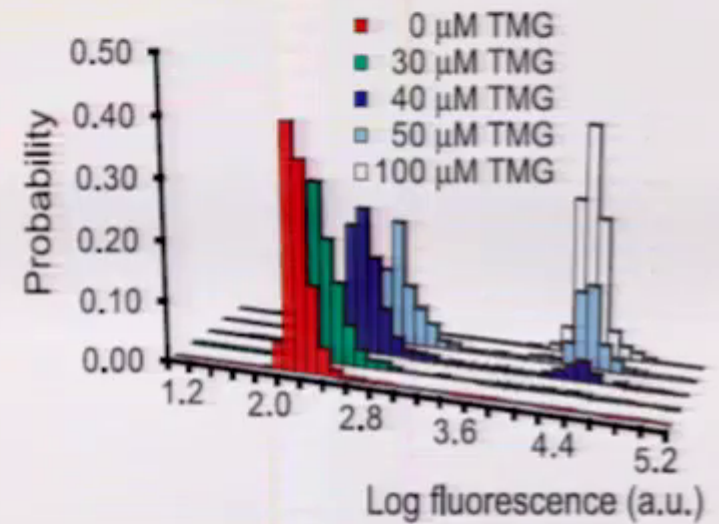


Relevant for bacterial persistence (x_0 as inducer or antibiotic)

Extrinsic noise induces bistability



Bimodal distributions



Choi *et. al.*, Science (2008)

Bistability range increases from $\sim 2 \mu\text{M}$ with no extrinsic noise to $\sim 30 \mu\text{M}$ for strong extrinsic noise