Modeling and control of gene regulatory networks

Madalena Chaves

BIOCO$_2$RE (Biological control of artificial ecosystems)

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
SOPHIA ANTIPOLIS - MÉDITERRANÉE

Math 640 Topics in control theory

Exercise 4.3.16 Consider a model for the "shopping cart" shown in Figure 4.2 ("knife-edge" or "unicycle" are other names for this example). The state is given by the orientation θ, together with the coordinates x_1, x_2 of the midpoint between the back wheels.

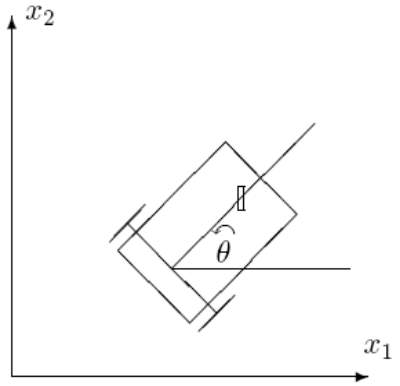


Figure 4.2: *Shopping cart.*

The front wheel is a castor, free to rotate. There is a non-slipping constraint on movement: the velocity $(\dot{x}_1, \dot{x}_2)'$ must be parallel to the vector $(\cos\theta, \sin\theta)'$. This leads to the following equations:

$$\begin{aligned} \dot{x}_1 &= u_1 \cos\theta \\ \dot{x}_2 &= u_1 \sin\theta \\ \dot{\theta} &= u_2 \end{aligned}$$

where we may view u_1 as a "drive" command and u_2 as a steering control (in practice, we implement these controls by means of differential forces on the two back corners of the cart). We view the system as having state space $\mathbb{R}^3$ (a more accurate state space would be the manifold $\mathbb{R}^2 \times \mathbb{S}^1$).
(a) Show that the system is completely controllable.
(b) Consider these new variables: $z_1 := \theta$, $z_2 := x_1 \cos\theta + x_2 \sin\theta$, $z_3 := x_1 \sin\theta - x_2 \cos\theta$, $v_1 := u_2$, and $v_2 := u_1 - u_2 z_3$. (Such a change of variables is called a "feedback transformation".) Write the system in these variables, as $\dot{z} = \widetilde{f}(z, v)$. Note that this is one of the systems Σ_i in Exercise 4.3.14. Explain why controllability can then be deduced from what you already concluded in that previous exercise. □

$\Rightarrow$ ARC holds at every $x^0 \in \mathbb{R}^3$.

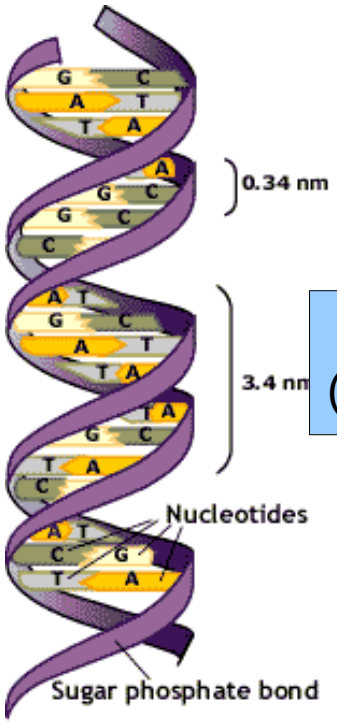
Therefore, by corollary 01.28, the shopping cart is completely controllable.

(something I'm glad to know, the next time I go to the supermarket!) (me too - can you imagine not being able to go from the entrance to the ice-cream aisle?)

(me too - can you imagine not being able to go from the entrance to the ice-cream aisle?)

Genetic networks: transcription and translation

DNA
(1-2 copy /cell)

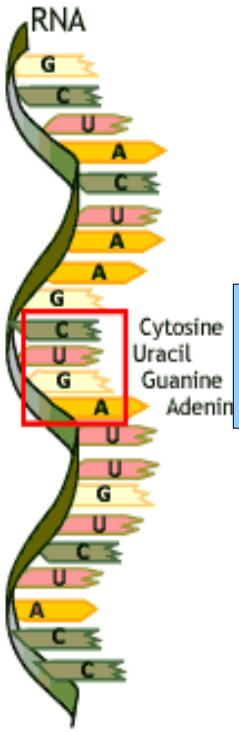


Transcription
(RNApolymerase)

1 min to transcribe

10^3 polymerase/cell

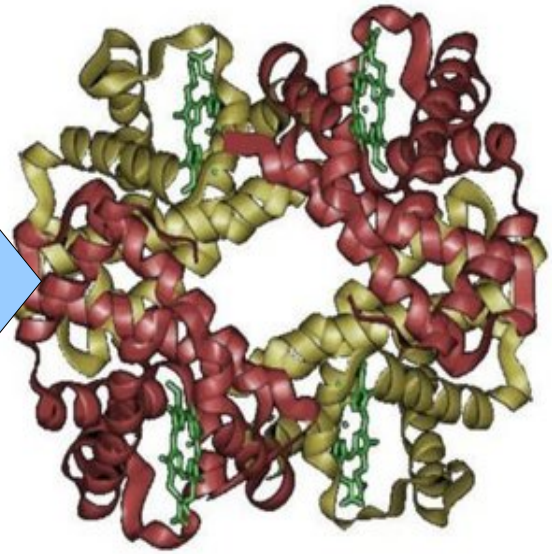
mRNA
(10^3 in E. coli)



2-5 min
mRNA lifetime

Translation
(Ribosomes)

Protein
(10^6 in E. coli
10^9 mammalian)

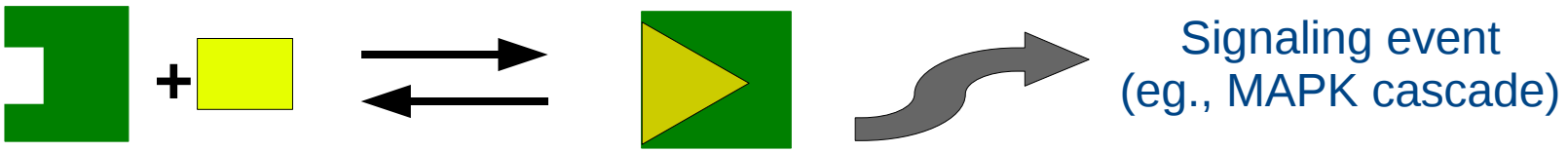


2 min to translate

10^4 ribosomes/cell

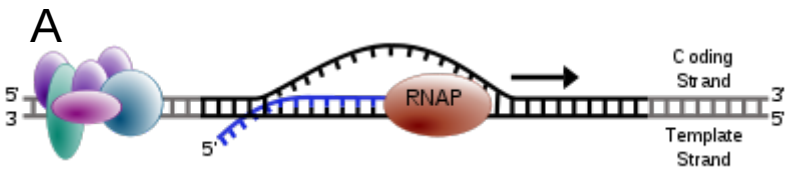
Genetic networks: some common interactions

Binding event

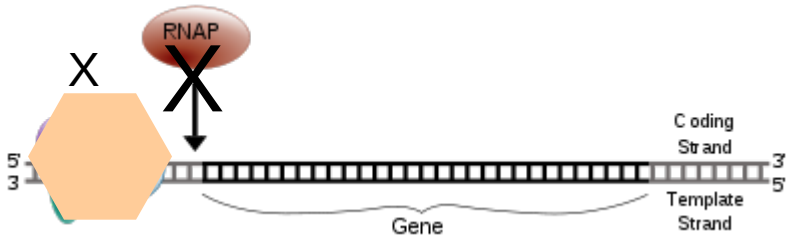


Signaling event
(eg., MAPK cascade)

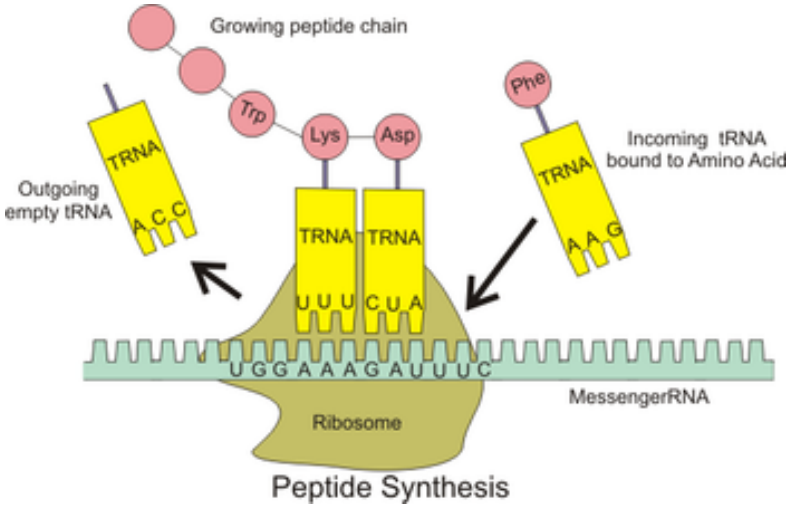
Activation of transcription (A → M)



Repression (X ⊣ M)



Translation (M → P)

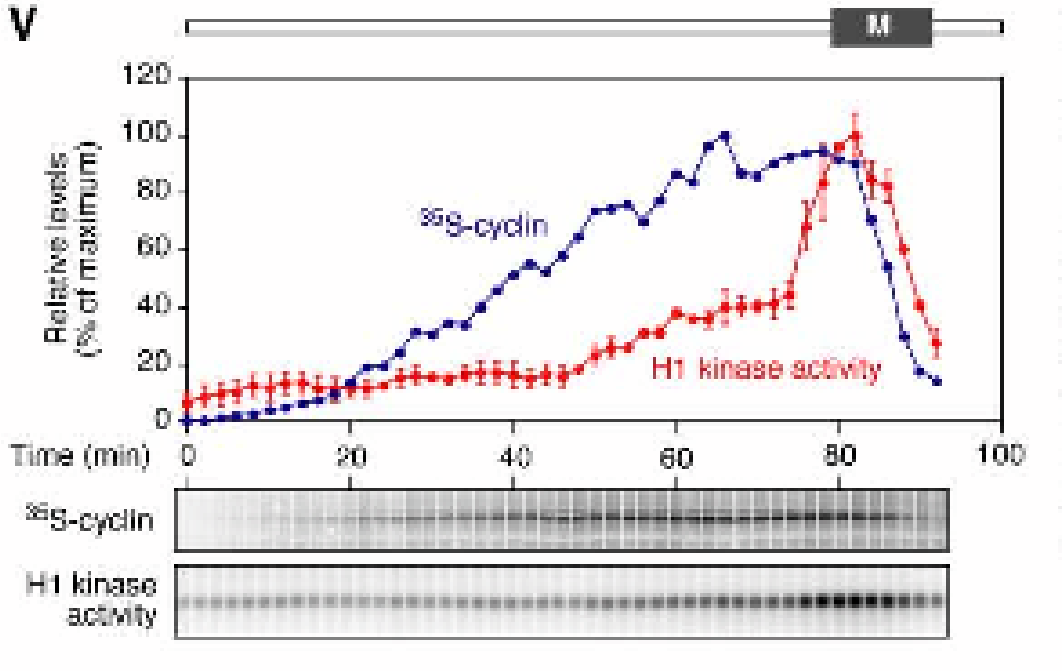


Experimental data ("data rich/data poor" Sontag 2005)

Expression of gene *wingless*, fly embryo (dark: higly expressed)

Microarray relative changes (red: expression increased)

Cdc2, cyclin B, Pomerening, Kim & Ferrell, Cell 2005



Genetic networks: questions and challenges

- Modeling

  Understanding the system; dynamics; predictions

- Model and experiments: available data

  different mathematical formalisms give different information

- Parameters

  calibration of models; robustness

Genetic networks: questions and challenges

- (Too) many components: model reduction techniques

  Two well-known modules: interconnection of two systems

- Control

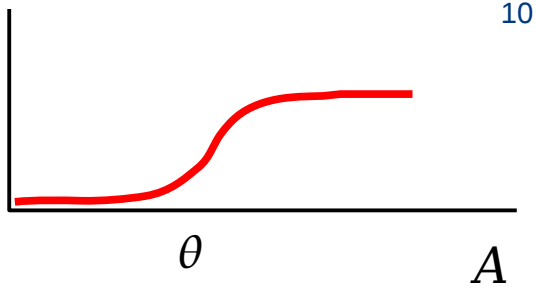
  How to find feedback laws?

  How to implement?

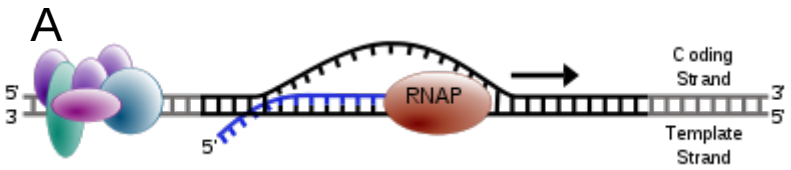
  Synthetic biology: assembling components; re-wiring a network

- State estimation, observers

Genetic networks: how to model



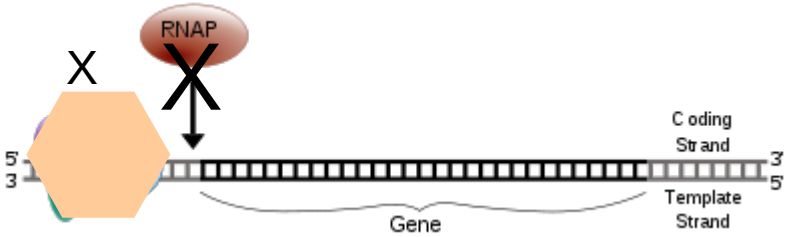
Activation of transcription (A ⟶ M)



Concentration of mRNA in terms of activator

$$\frac{dM}{dt} = \alpha \frac{A^n}{\theta^n + A^n} - \gamma_M M$$

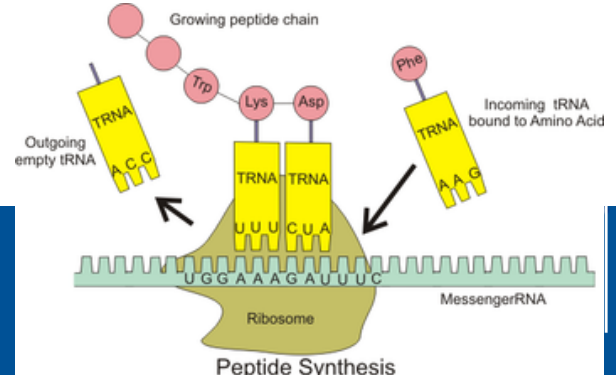
Repression (X ⟶| M)



Concentration of mRNA in terms of repressor

$$\frac{dM}{dt} = \alpha \frac{\theta^n}{\theta^n + X^n} - \gamma_M M$$

Translation (M ⟶ P)



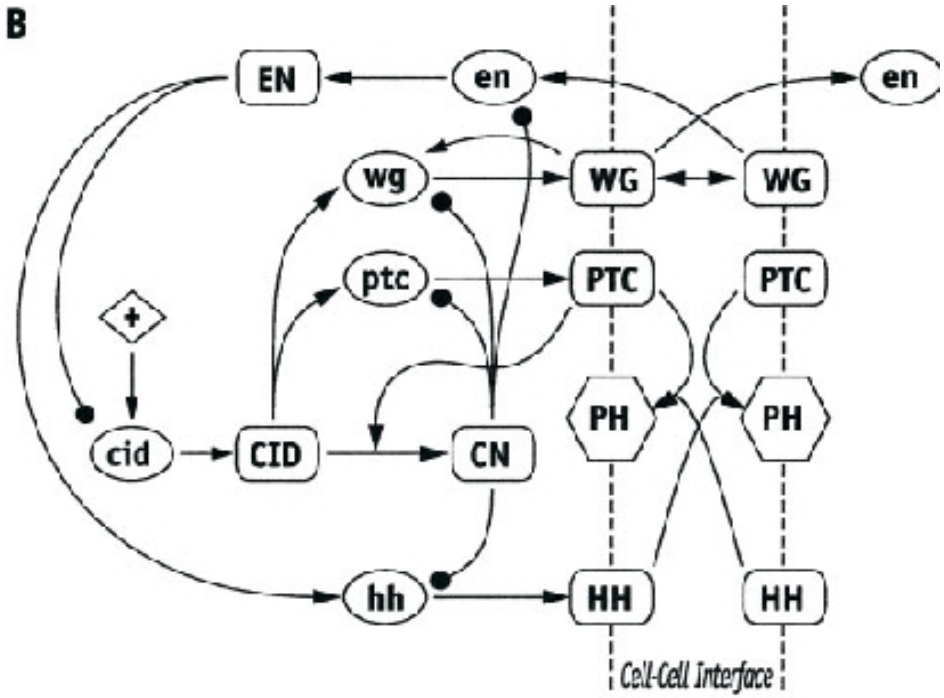
Concentration of protein

$$\frac{dP}{dt} = \alpha M - \gamma_P P$$

Example: drosophila segment polarity network

Model: **concentrations of mRNA and proteins**, for a group of 5 genes responsible for generating and maintaining the segmented body of the fruit fly

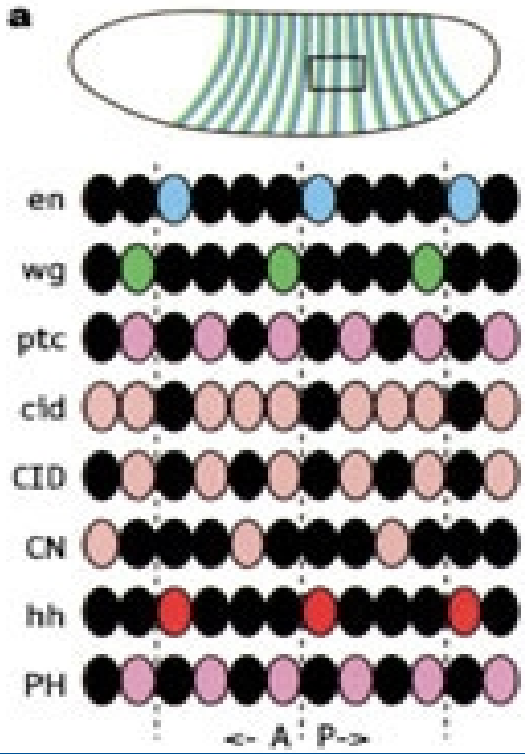
Goal: **reproduce** the observed **pattern of expression** for these 5 genes



Expression
of gene *wingless*

A model using ordinary differential equations

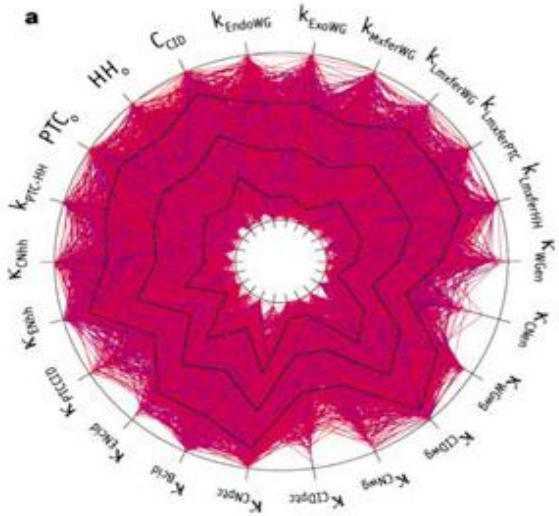
Drosophila segment polarity genes
von Dassow et al, Nature 2000



Parameters and dynamical behavior

Drosophila segment polarity genes
von Dassow et al, Nature 2000

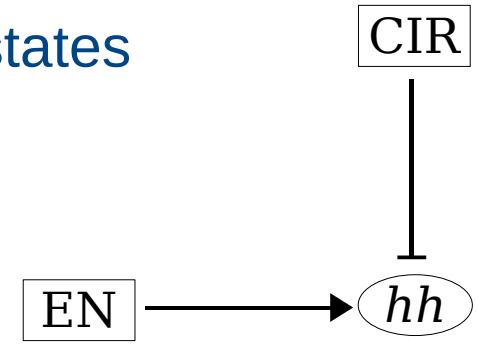
About 180 eqs.
Randomly try 200,000
sets of parameters
About 0.5% yield
"correct" gene pattern



Alternative frameworks: qualitative models

Boolean models: logical rules; 0/1 or ON/OFF states

$$hh(k+1) = \mathrm{EN}(k) \text{ and not } \mathrm{CIR}(k)$$



Robustness of the model **to perturbations** in the environment?

Fluctuations in the mRNA/protein concentrations;

Different timescales in biological phenomena;

Degradation and synthesis rates

A Boolean model of the segment polarity network

$$SLP_i^+ = \begin{cases} 0, & \text{if } i \in \{1,2\} \\ 1, & \text{if } i \in \{3,4\} \end{cases}$$

$$wg_i^+ = (CIA_i \text{ and } SLP_i \text{ and not } CIR_i) \text{ or } [wg_i \text{ and } (CIA_i \text{ or } SLP_i) \text{ and not } CIR_i]$$
$$WG_i^+ = wg_i$$

$$en_i^+ = (WG_{i-1} \text{ or } WG_{i+1}) \text{ and not } SLP_i$$
$$EN_i^+ = en_i$$

Albert & Othmer
J Theor Biol 2003

$$hh_i^+ = EN_i \text{ and not } CIR_i$$
$$HH_i^+ = hh_i$$

$$ptc_i^+ = CIA_i \text{ and not } EN_i \text{ and not } CIR_i$$
$$PTC_i^+ = ptc_i \text{ or } (PTC_i \text{ and not } HH_{i-1} \text{ and not } HH_{i+1})$$

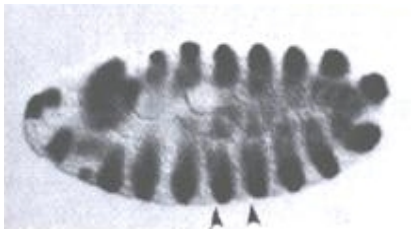
$$ci_i^+ = \text{not } EN_i$$
$$CI_i^+ = ci_i$$
$$CIA_i^+ = CI_i \text{ and } [\text{not } PTC_i \text{ or } HH_{i-1} \text{ or } HH_{i+1} \text{ or } hh_{i-1} \text{ or } hh_{i+1}]$$
$$CIR_i^+ = CI_i \text{ and } PTC_i \text{ and not } HH_{i-1} \text{ and not } HH_{i+1} \text{ and not } hh_{i-1} \text{ and not } hh_{i+1}$$

The model exhibits multiple “biological” equilibria

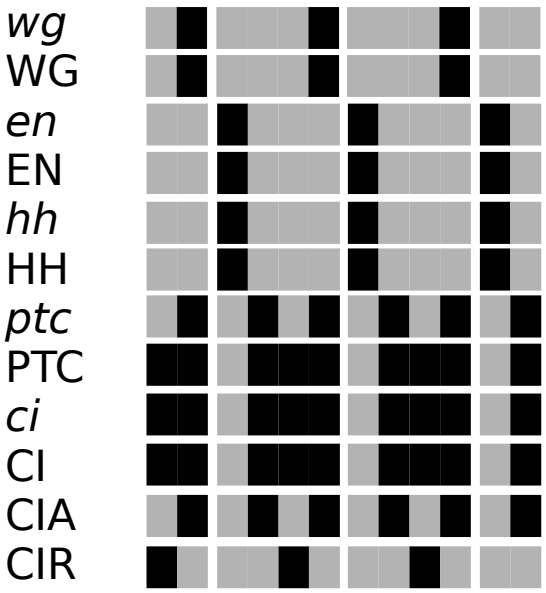
Wild type

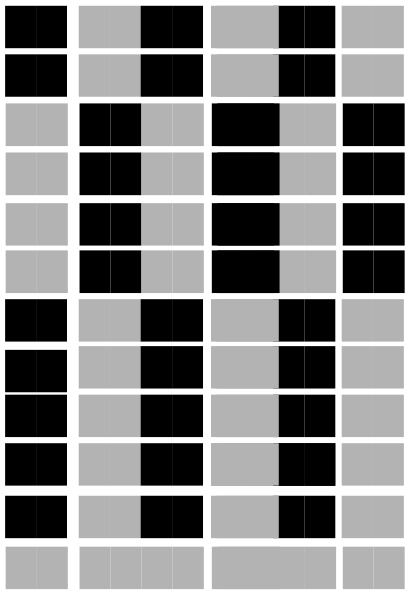
Broad stripes

No segmentation

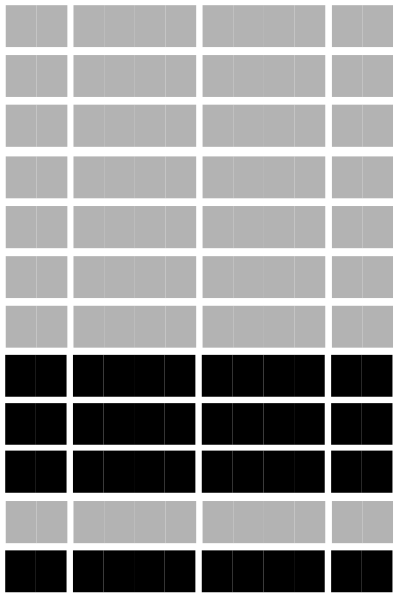


wg expression





ptc mutants,
heat shocked genes



en mutants
(lethal phenotype)

How to study Boolean models?

- Dynamics: synchronous or asynchronous algorithms?

$$hh(T_{hh}^{k+1}) = \text{EN}(T_{EN}^{k}) \text{ and not } \text{CIR}(T_{CIR}^{k})$$

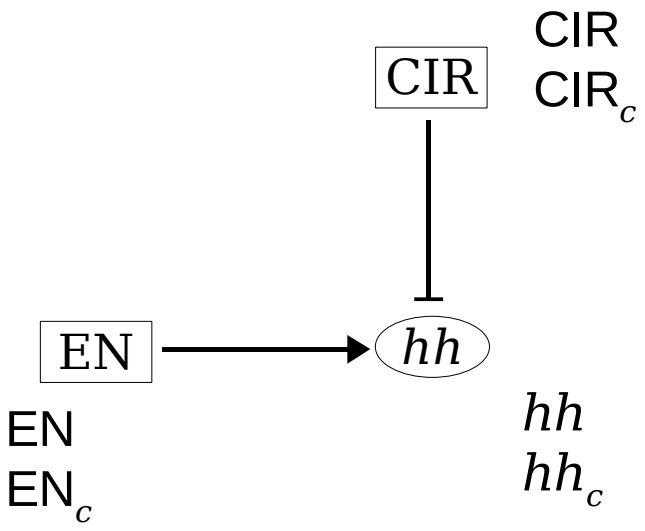
Chaves, Albert & Sontag, JTB 2005

- Piecewise linear models - Glass type

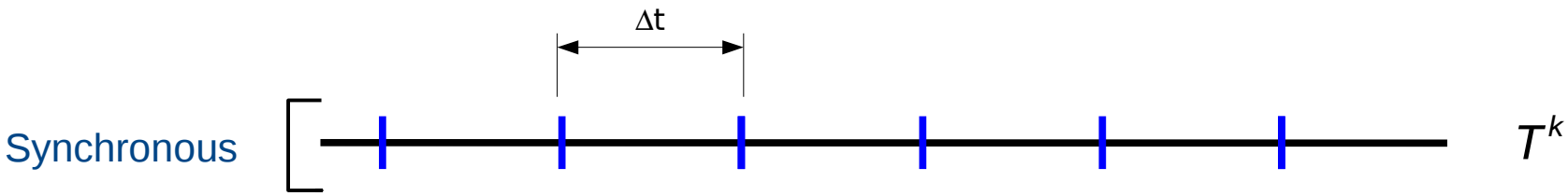
$$\frac{dhh_c}{dt} = -\alpha_i \ (hh_c + F_{hh})$$

with: $F_{hh}(t) = \text{EN}(t) \text{ and not } \text{CIR}(t)$

$$hh = \begin{cases} 0, & \text{if } hh_c < 0.5 \\ 1, & \text{if } hh_c > 0.5 \end{cases}$$

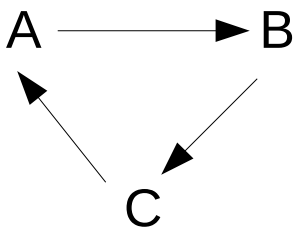


Boolean models: updates and dynamics

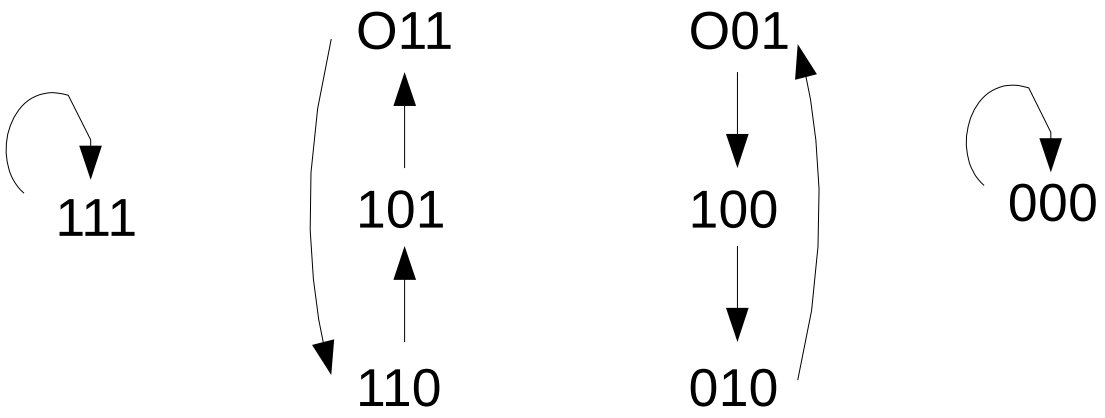


⇒ All variables simultaneously updated.
Deterministic trajectories in a directed graph.

Positive loop

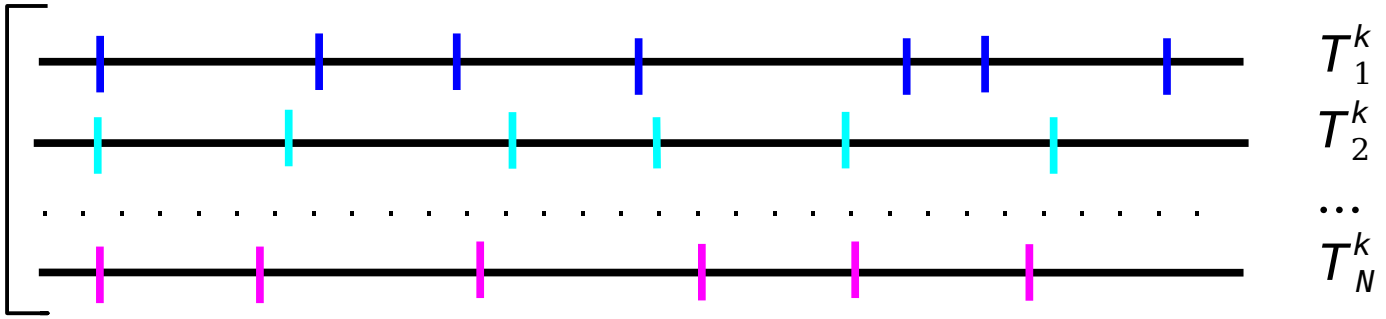


Synchronous transition graph



Boolean models: updates and dynamics

Asynchronous



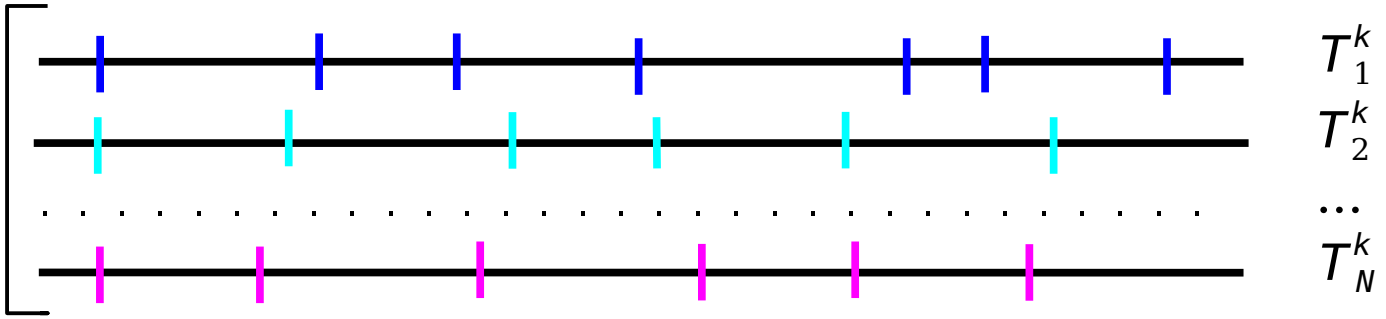
Each variable updated at its own pace:

perturbed time unit $(1+ r)\ T$, r in $[-\varepsilon, \varepsilon]$

$\Rightarrow$ NOT deterministic

Boolean models: updates and dynamics

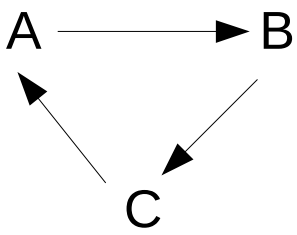
Asynchronous

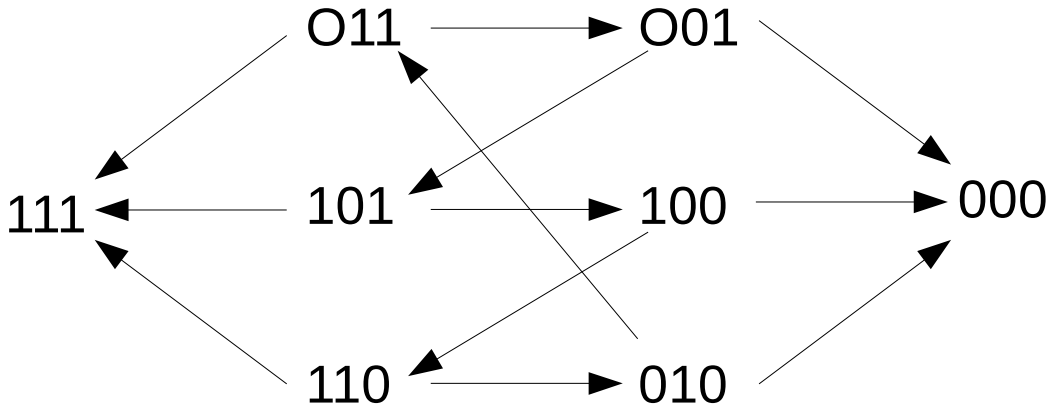


Each variable updated at its own pace:

perturbed time unit (1+ r) T , r in [-ε, ε]

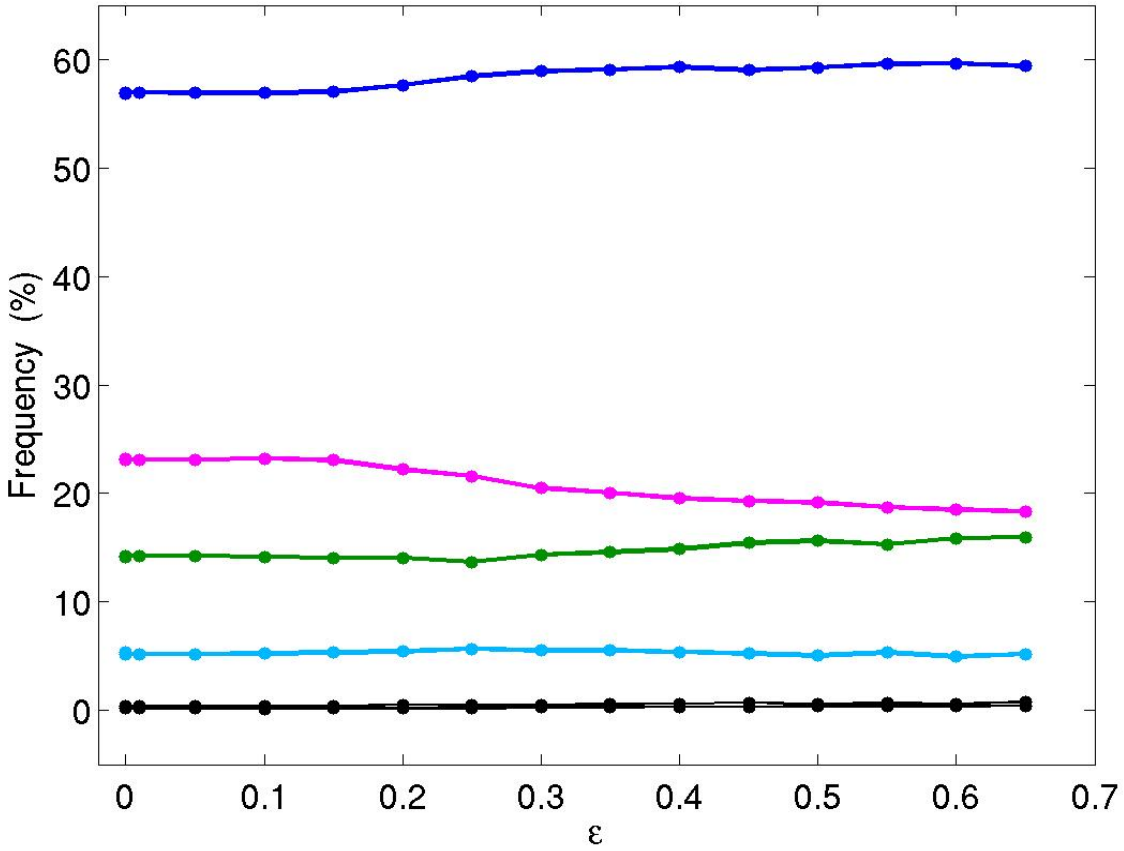
Follow one of many possible trajectories in the asynchronous transition graph,





Totally asynchronous and random order updates

Starting from same initial state, percentage of simulations that converge to each steady state ----- **low robustness**...



56%, Wild type

24%, Broad stripes

15%, No segmentation

4%, Wild type variant

1%, Ectopic and variant

Random order updates + Timescale separation

First, update all protein nodes; then, update all mRNA nodes
Any permutation among protein nodes followed by any permutation among mRNA nodes

Theorem: Trajectories diverge from the wild type steady state if and only if the first permutation among proteins satisfies the following order, in the third cell

CIR_3	$\mathbf{CI_3}$	$\mathbf{CIA_3}$	$\mathbf{PTC_3}$	
$\mathbf{CI_3}$	CIR_3	$\mathbf{CIA_3}$	$\mathbf{PTC_3}$	[CI-PTC]
$\mathbf{CI_3}$		$\mathbf{CIA_3}$	CIR_3 $\mathbf{PTC_3}$	

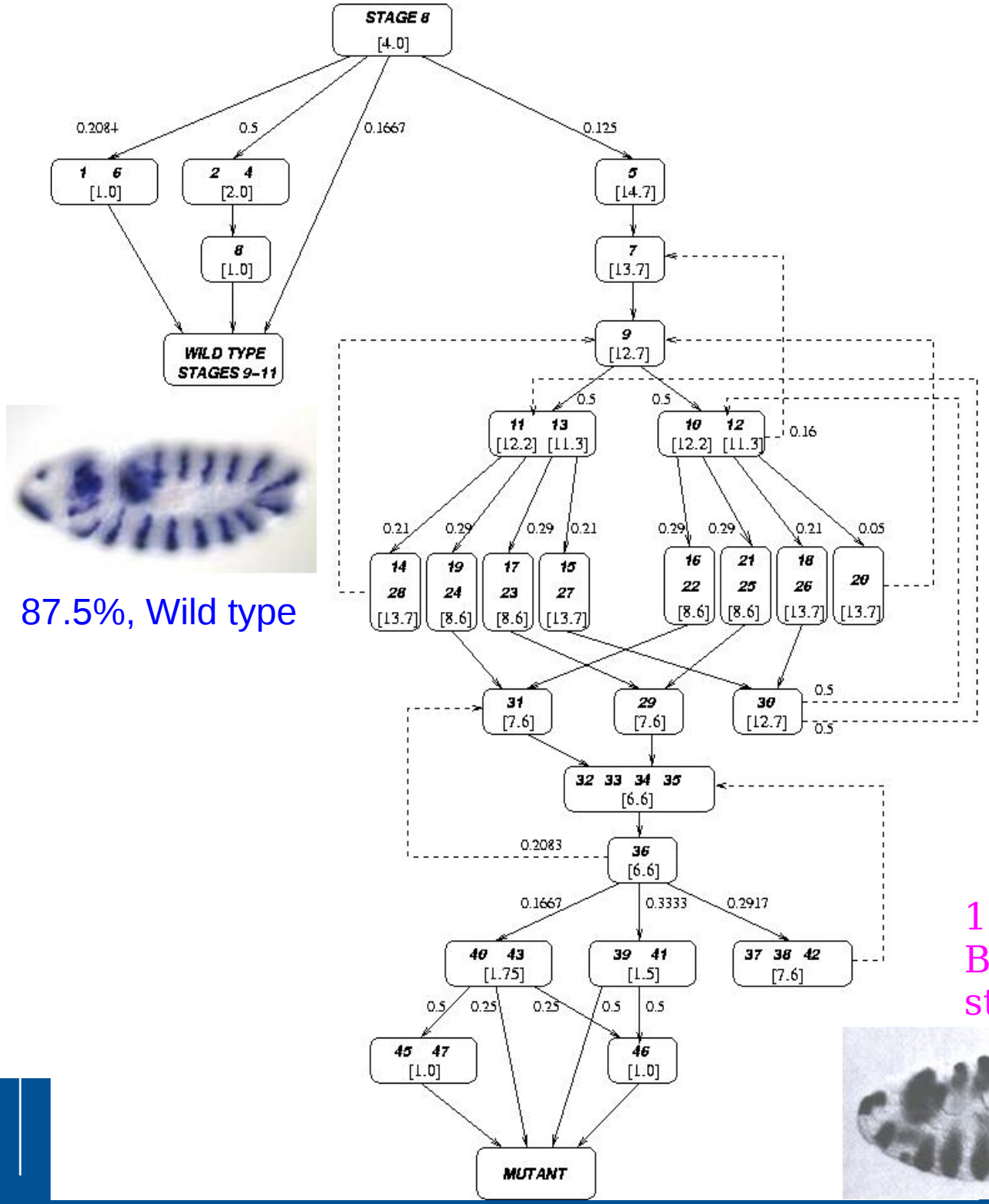
and all other proteins may appear in any of the remaining sites.

Chaves, Albert & Sontag, JTB 2005

Random order
+
Timescale separation

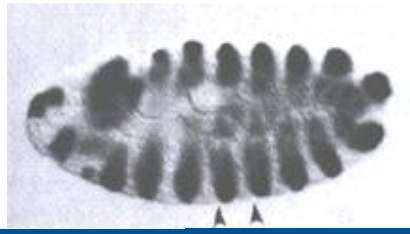
Markov Chain with two absorbing states

Increased robustness



87.5%, Wild type

12.5%, Broad stripes



Piecewise linear systems: Glass-type model

Timescale of node X_i

Synthesis of gene/protein X_i (ON/OFF)

$$\frac{dx_i}{dt} = \alpha_i \left(F_i(X_1,\ldots,X_n) - x_i\right)$$

with: $F_i(X_1,\ldots,X_n)$ = Boolean rule for node X_i

Steady states: same as in Boolean model

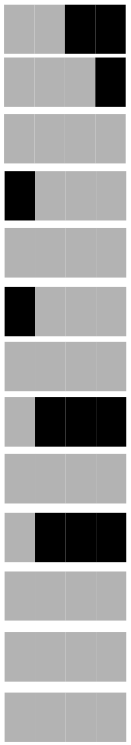
and:
$$X_i = \begin{cases} 0, & \text{if } x_i < \theta_i \\ 1, & \text{if } x_i > \theta_i \end{cases}$$

Based on: Glass& Kauffman, 1973; Edwards and Glass, 2000

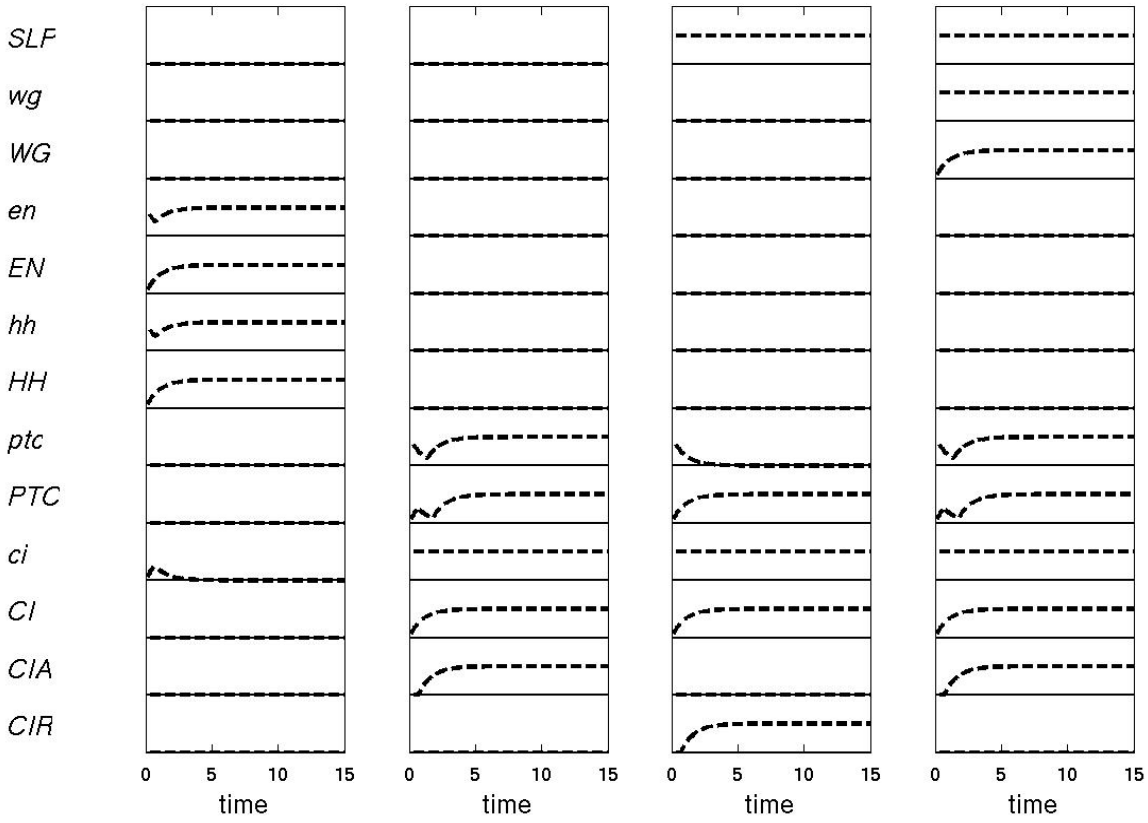
Some simulations

Four cells in each parasegment; periodic boundary conditions

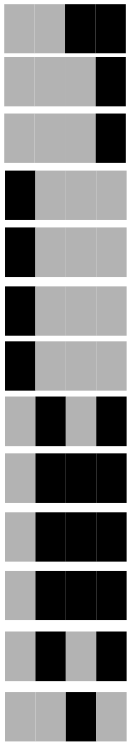
Initial (stage 8)



Cell 1 Cell 2 Cell 3 Cell 4



Final (stages 9-11)



Timescale separation: 100% convergence to WT

Assumption I: $\alpha_{\text{protein}} > 2\,\alpha_{mRNA}$

Assumption II: $\theta_1 = \theta_i \leq 0.5$

Assumption III: $\alpha_{\text{PTC}_3} > \alpha_{\text{CI}_3}$

Theorem: Under these assumptions the Glass-type model always converges to the wild type steady state

Chaves, Sontag & Albert 2006

Analysis of Boolean models and beyond

- Robustness and fragility of Boolean models for genetic regulatory networks, **Chaves, Albert and Sontag, 2005**:
Paper was **in JTB top 10 most cited (of the last 5 years)**

- "Timescale separation" leads to "Priority classes"
(Bioinformatics: GINsim software Chaouiya, Thieffry, etc.)

- Further work: asynchronous transition graphs and the **dynamical behavior of "large" networks**

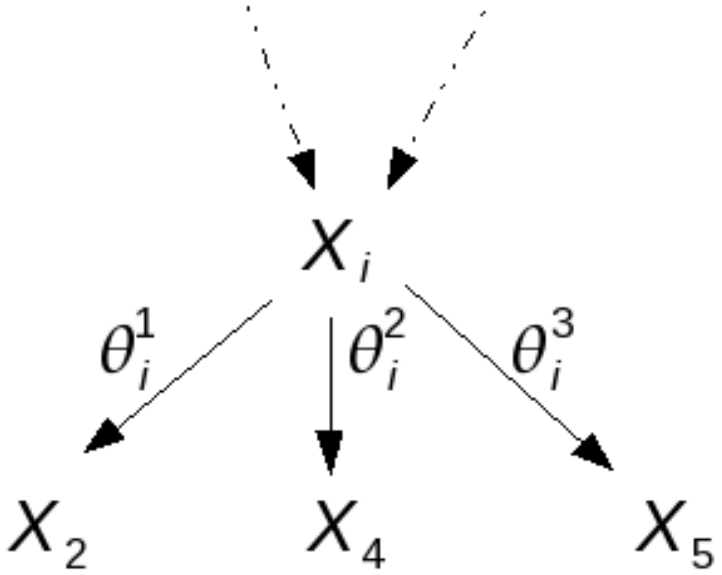
- Further work: **piecewise linear systems**

Piecewise linear systems: qualitative framework

$$\dot{x} = f(x) - \Gamma x$$

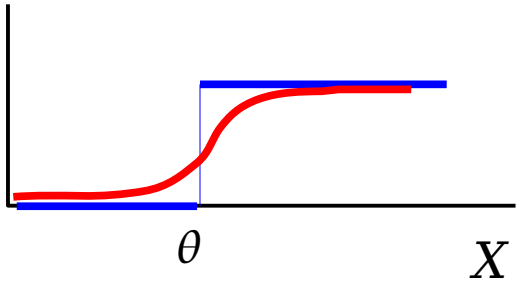
$$x \in \mathbb{R}^n_{\geq 0}, \quad f : \mathbb{R}^n_{\geq 0} \times \mathbb{R}^n_{\geq 0}, \quad \Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$$

Thresholds: $0 < \theta_i^1 < \cdots < \theta_i^{r_i} < M_i$



Function *f* is a sum of products of step functions

$s^+(X, \theta)$



Refs: Casey, de Jong & Gouzé, 2006

Piecewise linear systems

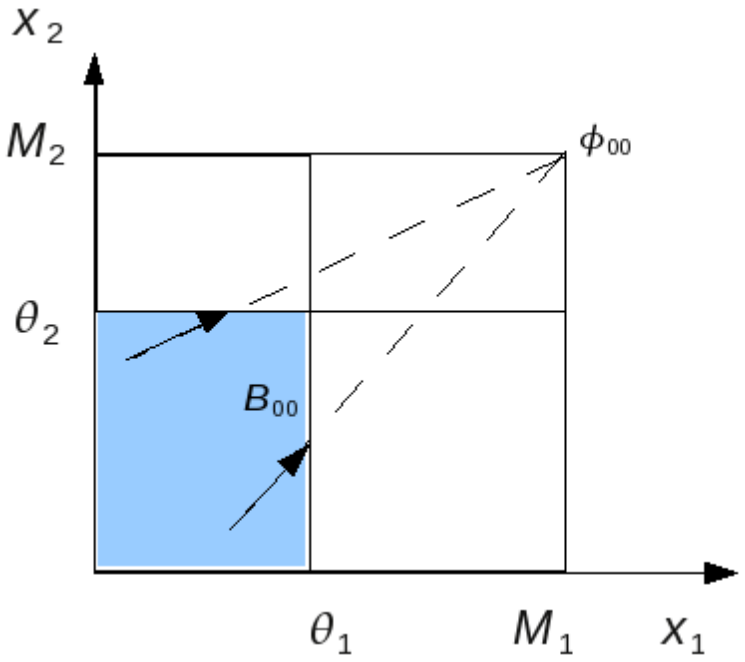
Regular domains: $B_{k_1,\ldots,k_n}, \quad k_i \in \{0, r_i\}, \quad \theta_i^{k_i} < x_i < \theta_i^{k_i+1}$

Switching domains: $D_l, \quad x_i = \theta_i^l, \quad$ for some i

Focal points: $\dot{x} = f^{k_1,\ldots,k_n} - \Gamma\, x = 0 \;\Rightarrow\; \phi^{k_1,\ldots,k_n} = \Gamma^{-1} f^{k_1,\ldots,k_n}$

Example:

$$\dot{x}_1 = \kappa_1 \; s^-(x_2, \theta_2) - \gamma_1 x_1$$
$$\dot{x}_2 = \kappa_2 \; s^-(x_1, \theta_1) - \gamma_2 x_2$$



Measurements and control

Qualitative measurements:

$$s^+(x_i, \theta_i^r) \in \{0, 1\}$$

Know only: position of variables with respect to thresholds
(either "**weakly expressed**" or "**strongly expressed**")

Qualitative inputs: u piecewise constant (in each regular domain)

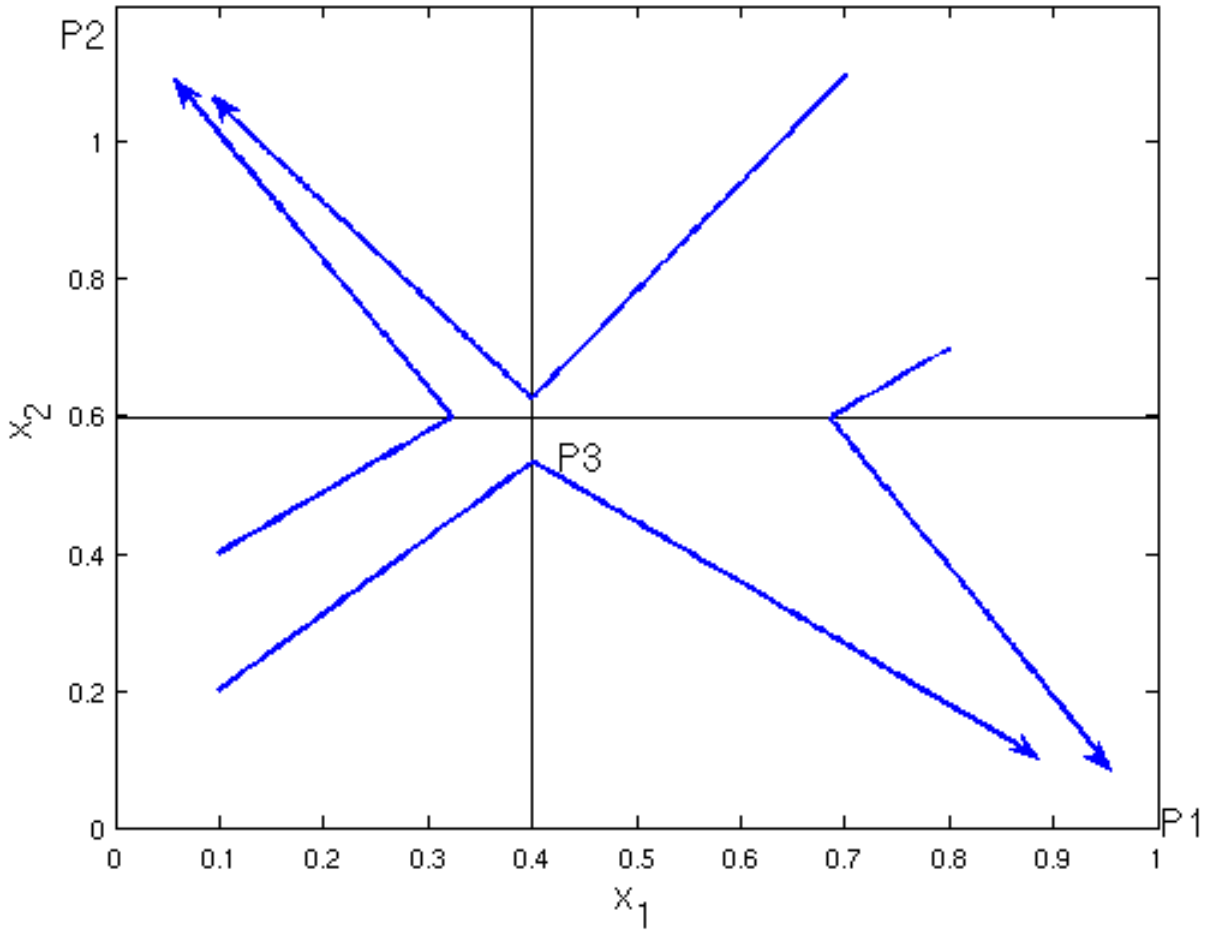
$$u: \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}^n \rightarrow \{u_{\min}, 1, u_{\max}\}$$

Can **only implement three values**.
Inputs can act on degradation or synthesis rates (inducers)

Chaves & Gouzé, Automatica 2011

Control of simple biological motifs: the bistable switch

$$\dot{x}_1 = \kappa_1 \; s^-(x_2, \theta_2) - \gamma_1 x_1 \;, \quad \dot{x}_2 = \kappa_2 \; s^-(x_1, \theta_1) - \gamma_2 x_2$$

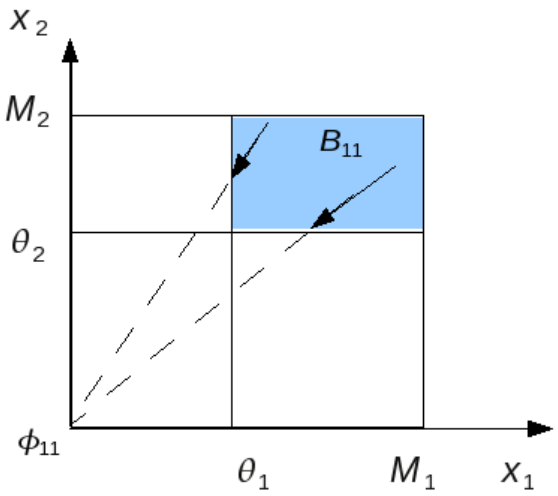


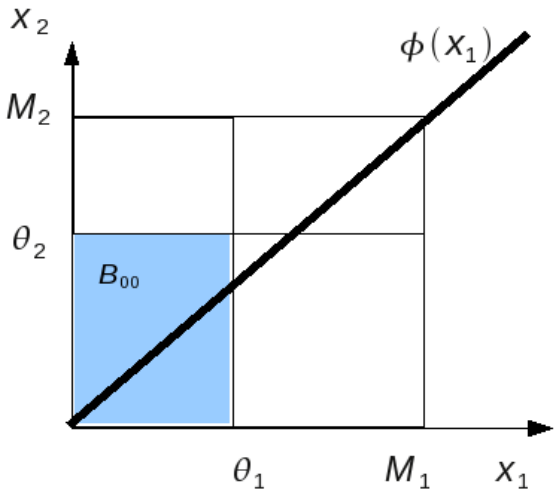
Control of the bistable switch

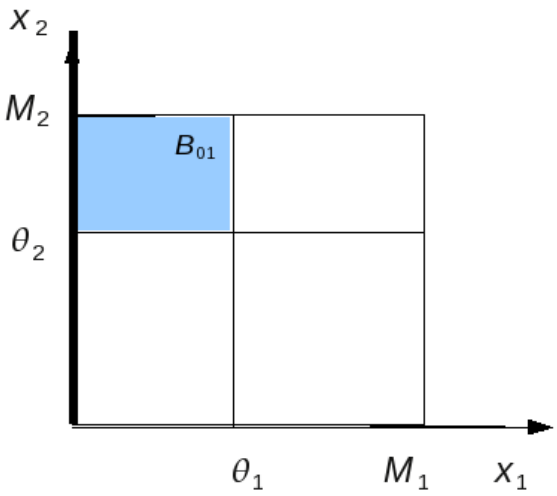
$$\dot{x}_1 = \boldsymbol{u}\kappa_1 s^-(x_2,\theta_2) - \gamma_1 x_1 \ , \quad \dot{x}_2 = \boldsymbol{u}\kappa_2 s^-(x_1,\theta_1) - \gamma_2 x_2$$

Problem: using only qualitative control laws, is it possible to drive the system to either of its stable steady states?

Control: relocate focal points

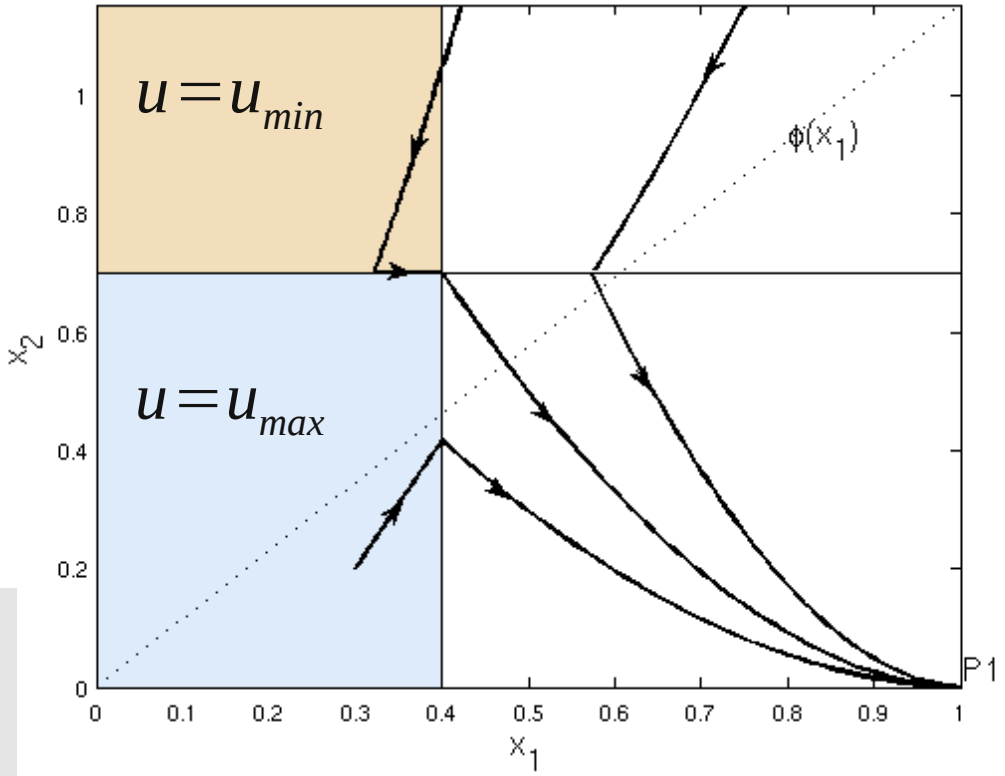






Control to steady state P1

$$u(x) = \begin{cases} 1 & x \in B_{11} \cup B_{10} \\ u_{\min} & x \in B_{01} \\ u_{\max} & x \in B_{00} \end{cases}$$

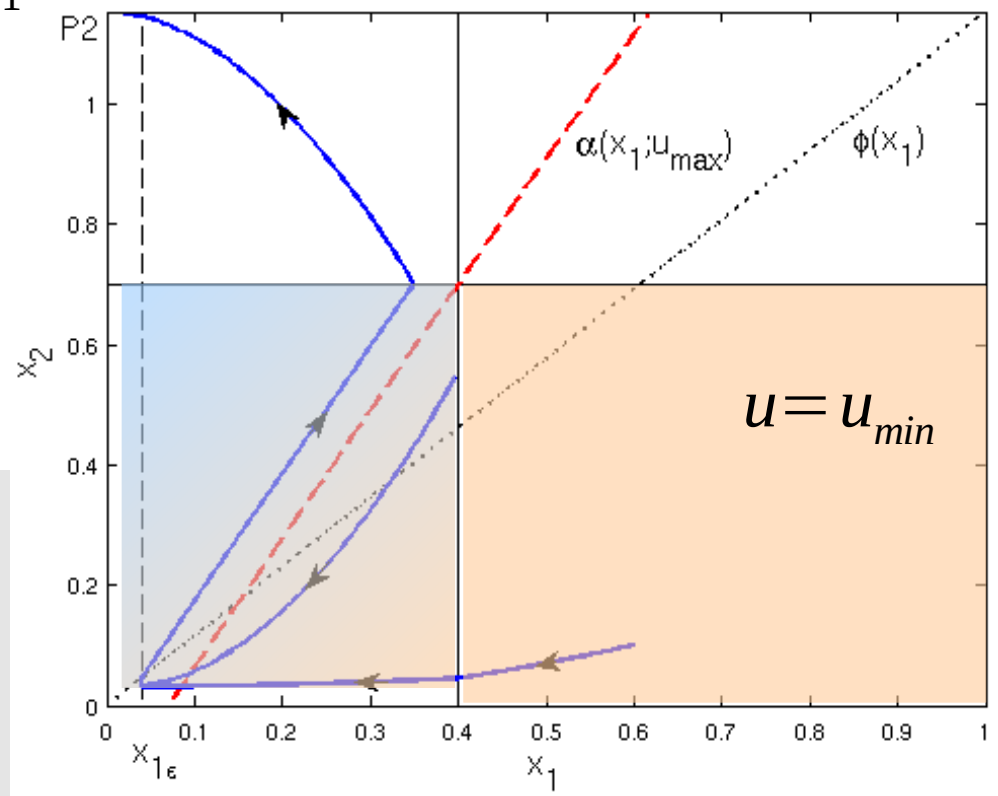


Theorem: Assume that $\Phi(\Theta_1)<\Theta_2$.
The system with this control law converges to point P1.

Chaves & Gouzé, Automatica 2011

Control to steady state P2

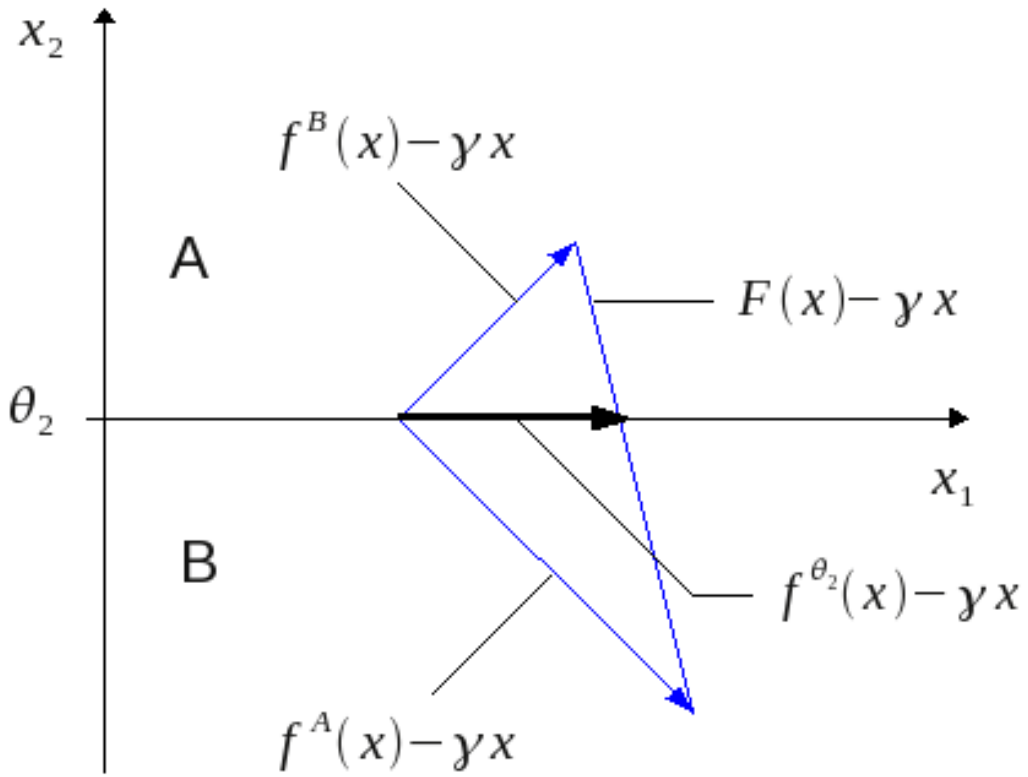
$$u(t,x) = \begin{cases} 1 & \forall\, t, x \in B_{11} \cup B_{01} \\ u_{\min} & \forall\, t, x \in B_{01} \\ u_{\min} & \forall\, t < T_1, x \in B_{00} \\ u_{\max} & \forall\, t \geq T_1, x \in B_{00} \end{cases}$$



Theorem: Assume that $\Phi(\Theta_1)<\Theta_2$, and condition on separatrix.
The system with this control law converges to point P2.

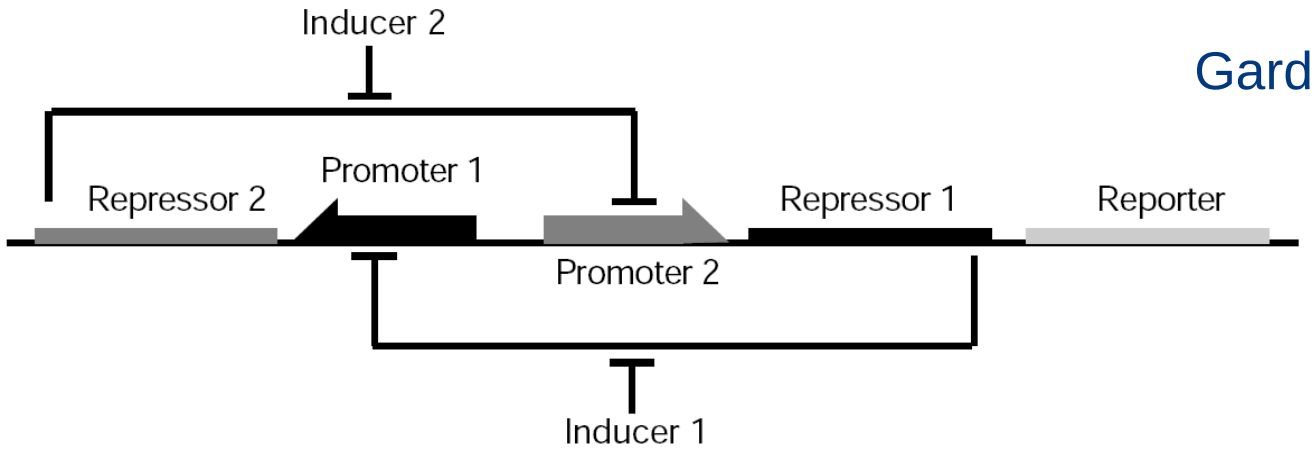
Chaves & Gouzé, Automatica 2011

Using Filippov solutions

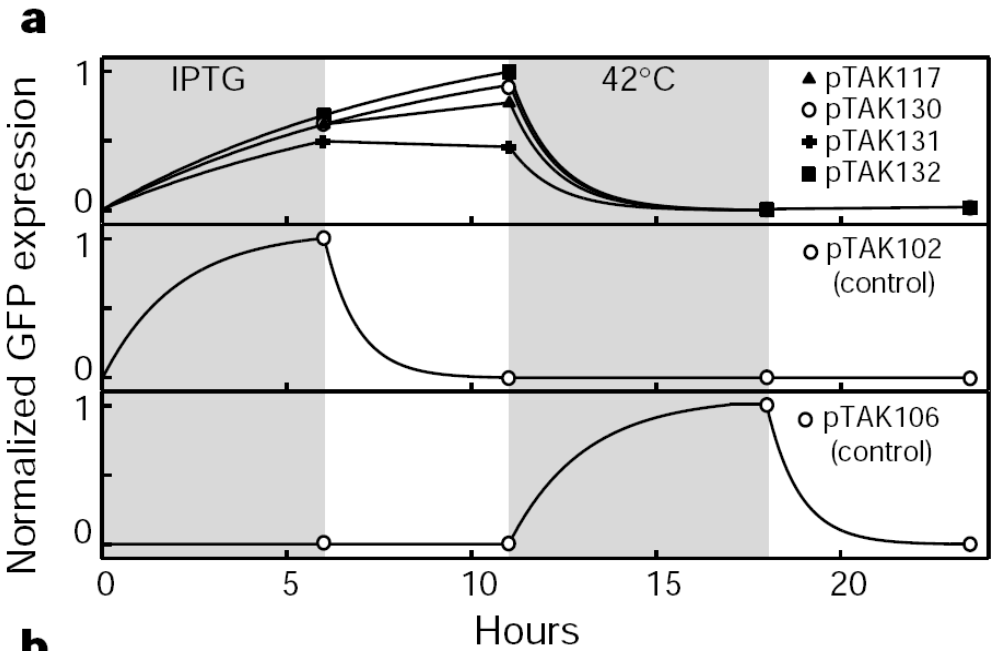


$$\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix} \in \overline{\text{co}}\{\ f^A(x)-\gamma x,\ f^B(x)-\gamma x\ \}$$

A synthetic bistable switch



Gardner, Cantor & Collins, Nature 2000



$$u \approx \frac{\text{IPTG}}{\text{Temperature}}$$

$$u_{\max} \approx \text{Apply IPTG}$$

$$u_{\min} \approx \text{High temperature}$$

Conclusions

- Experimental data: choose **appropriate formalism**
  different formalisms provide **complementary information**

- **Qualitative control**
  find feedback laws using only qualitative data (for simple motifs)

  easier to implement
  “add large amount of inducer when expression of X is high”

  synthetic biology: assembling components; re-wiring a network

- Boolean models:
  large networks as interconnection of two smaller modules

Final conclusion

THANK YOU EDUARDO

.... AND CONGRATULATIONS !