Validation of Qualitative Models of Genetic Regulatory Networks by Model Checking: Nutritional Stress Response in *E. coli*

Grégory Batt,[1] Delphine Ropers,[1] Hidde de Jong,[1] Johannes Geiselmann[2], Radu Mateescu,[1] Michel Page,[1] Dominique Schneider[2]

[1] INRIA Rhône-Alpes, Grenoble

[2] Laboratoire Adaptation et Pathogénie des Microorganismes, Université Joseph Fourier, Grenoble

Email: Gregory.Batt@inrialpes.fr

Overview

1. Introduction

2. Qualitative modeling and simulation method

3. Validation with model checking techniques

4. Application to nutritional stress response in *E. coli*

5. Conclusions

Genetic regulatory networks

❖ **Genetic regulatory networks** underlie functioning and development of living organisms

Genes, proteins, small molecules, and their regulatory interactions
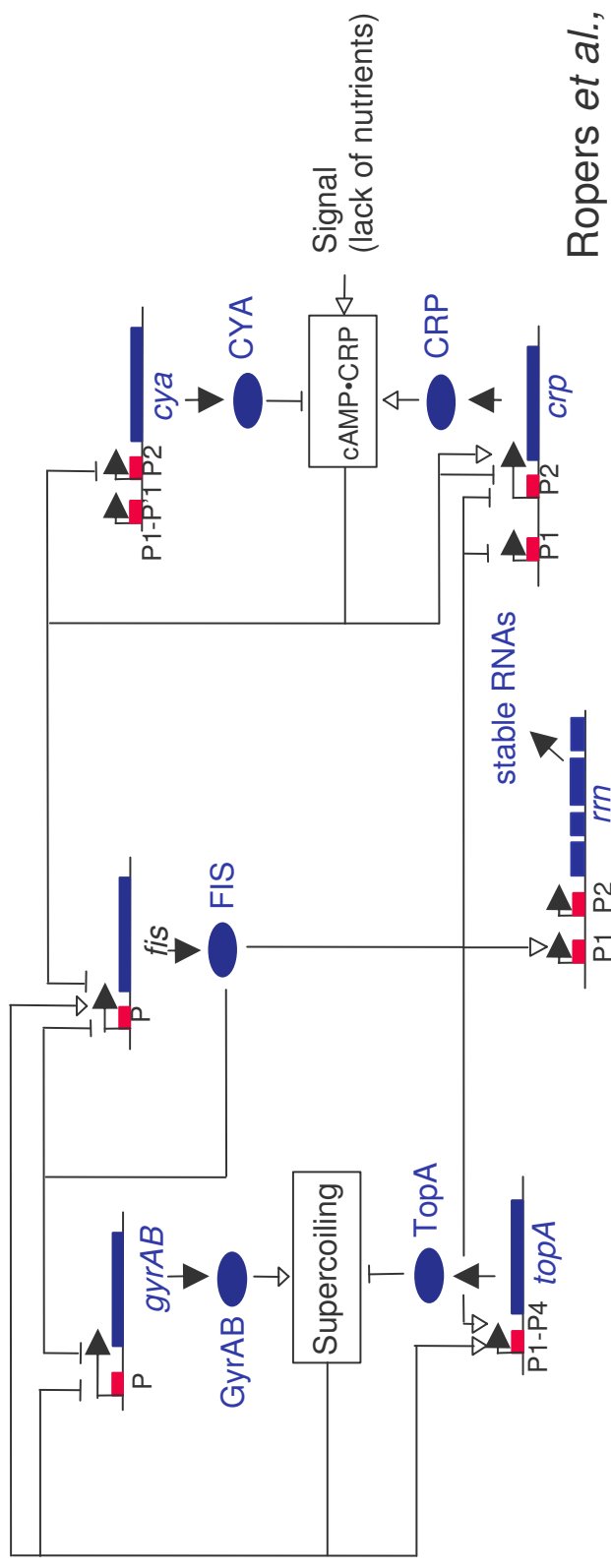
Ropers *et al.*, 05

❖ Genetic regulatory networks are **large** and **complex**

Validation of genetic network models

❖ Mathematical methods and computational tools exist for analysis of genetic networks through modeling and simulation

❖ Problem of **model validation**: check consistency between predictions and experimental data

❖ Major constraints for model validation:

- predictions **suitable for comparison** with available experimental data

- **automatic** and **efficient** comparison between predictions and experimental data

❖ Approach:

- refinement of qualitative modeling and simulation method
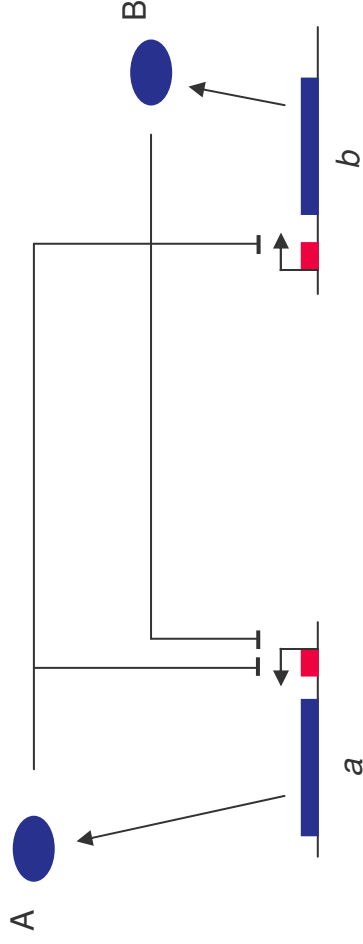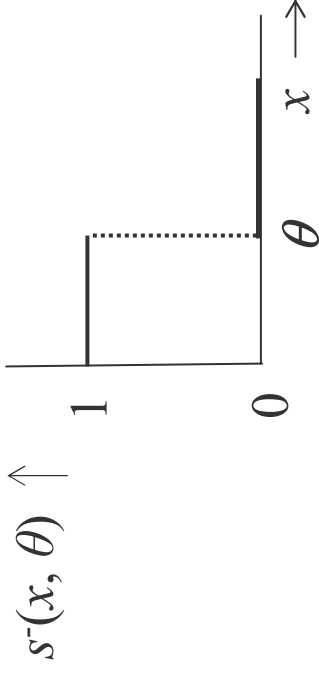
- combination with model checking techniques

PL differential equation models

❖ Genetic networks modeled by class of differential equations using **step functions** to describe regulatory interactions

$$\dot{x}_a = \kappa_a \, s^-(x_a, \, \theta_{a2}) \, s^-(x_b, \, \theta_b) - \gamma_a \, x_a$$

$$\dot{x}_b = \kappa_b \, s^-(x_a, \, \theta_{a1}) - \gamma_b \, x_b$$

x : protein concentration
θ : threshold concentration
κ, γ: rate constants

❖ Differential equation models of regulatory networks are **piecewise-linear (PL)**

Qualitative analysis of network dynamics

❖ Method for **qualitative analysis** of dynamics of genetic regulatory networks:

- Treatment of **discontinuities** in step functions by generalizing differential equations to differential inclusions (Filippov-like approach)

- Discrete transition system obtained by **qualitative abstraction** based on hyperrectangular partition of phase space

- **Inequality constraints** define regions in parameter space yielding the same discrete transition system

- **Symbolic computation** of transition system using inequality constraints and tailored algorithms (upscalability)

- Implementation in the computer tool **Genetic Network Analyzer (GNA)**

- **Application** to initiation of sporulation in *B. subtilis*

Gouzé and Sari, *Dyn. Syst.*, 03   de Jong et al., *Bull. Math. Biol.*, 04

Problem for model validation

❖ Model validation using gene expression data: observation of **changes in derivative signs**

$x_a > 0 \qquad \dot{x}_a < 0$

$\dot{x}_b > 0 \qquad \dot{x}_b < 0$
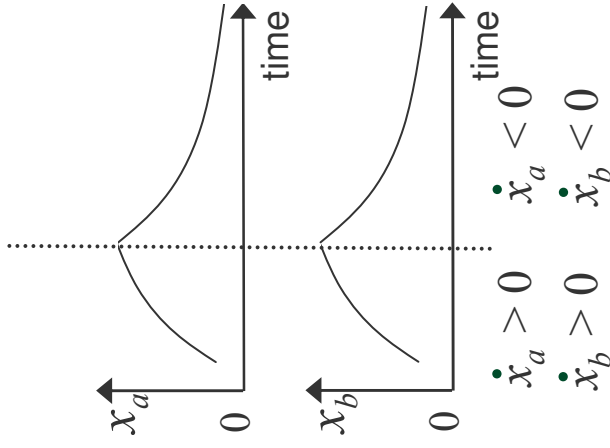
$$\dot{x}_a = \kappa_a \, s^-(x_a, \theta_{a2}) \, s^-(x_b, \theta_b) - \gamma_a \, x_a$$

$$\dot{x}_b = \kappa_b \, s^-(x_a, \theta_{a1}) - \gamma_b \, x_b$$

Problem for model validation

❖ Model validation using gene expression data: observation of **changes in derivative signs**

max_b

M^{11}	M^{12} M^{13}	M^{14} M^{15}
M^6	M^7 M^8	M^9 M^{10}
M^1	M^2 M^3	M^4 M^5

θ_b

0 θ_{a1} θ_{a2} max_a

x_a

0 time

x_b

0 time

- $\dot{x}_a > 0$ $\dot{x}_a < 0$
- $\dot{x}_b > 0$ $\dot{x}_b < 0$

UNIVERSITE JOSEPH FOURIER
SCIENCES TECHNOLOGIE MEDECINE
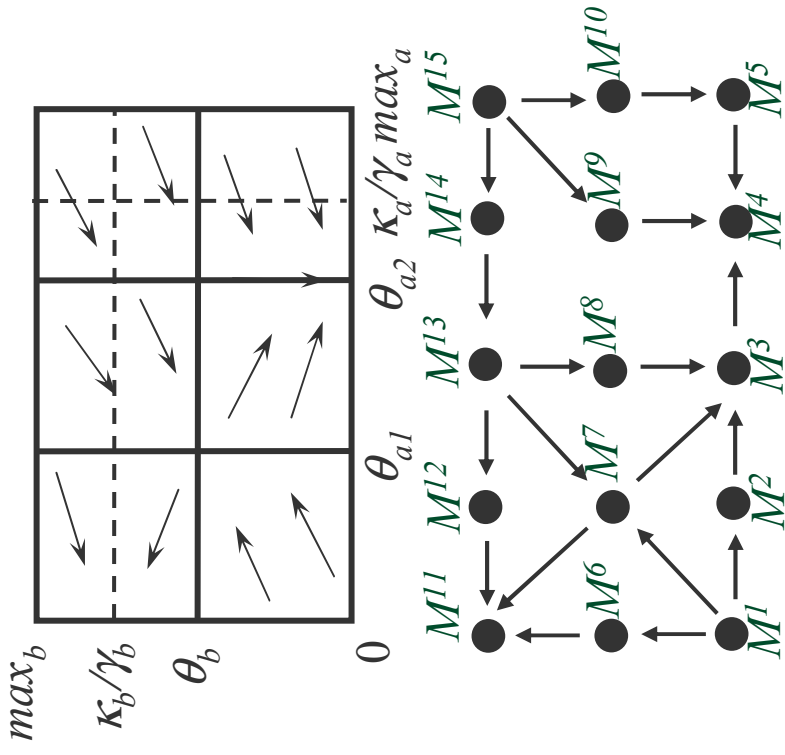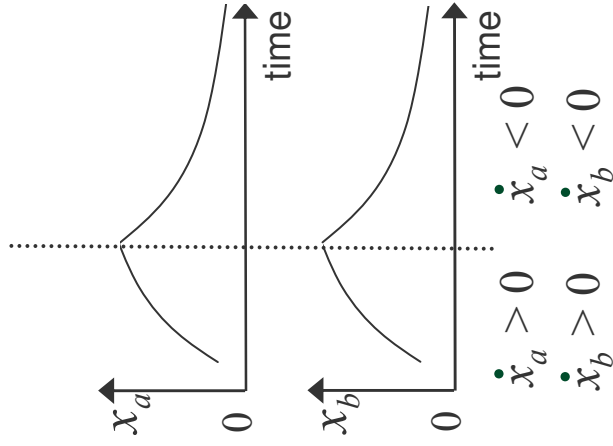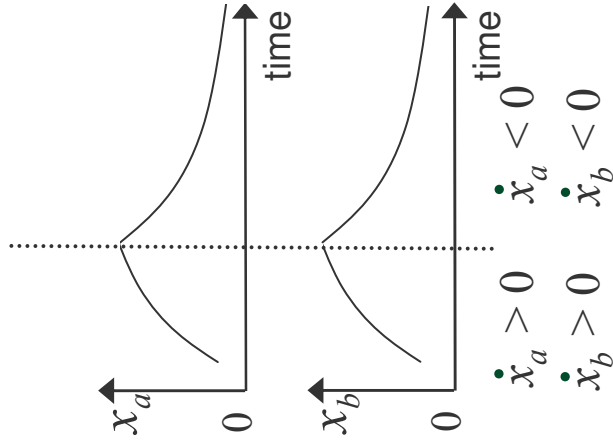
INRIA
RHÔNE-ALPES
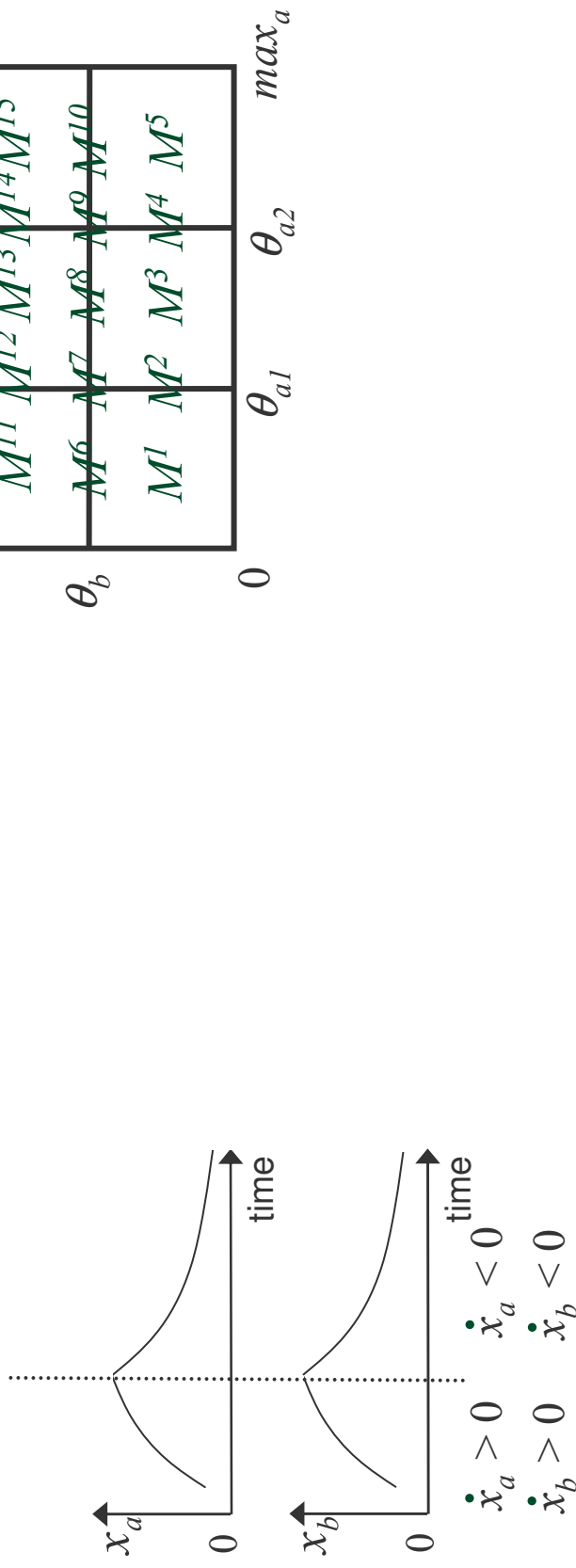
Problem for model validation

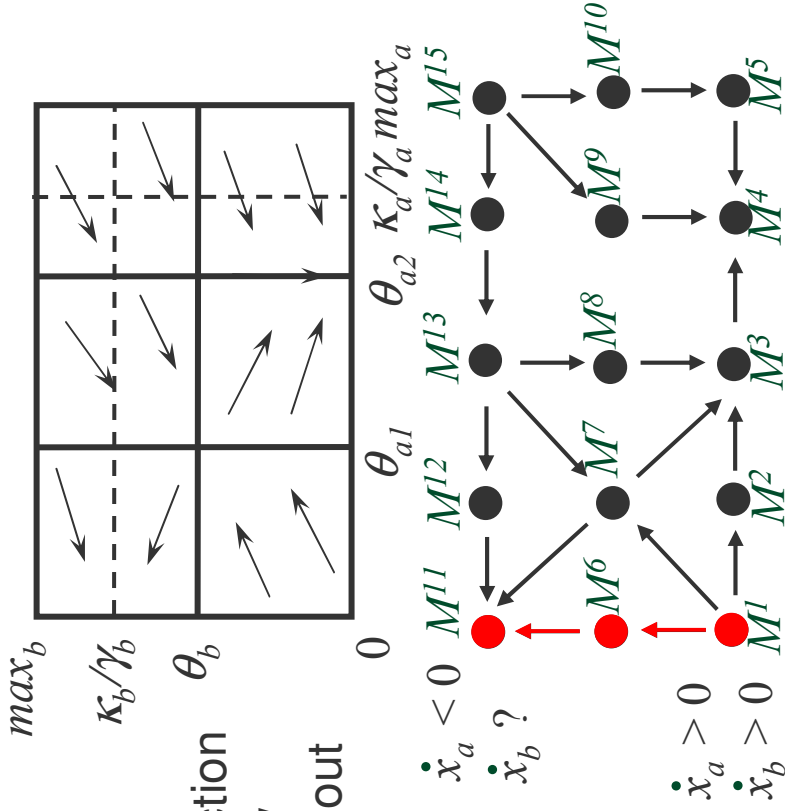❖ Model validation using gene expression data: observation of **changes in derivative signs**

Problem for model validation

❖ Model validation using gene expression data: observation of **changes in derivative signs**

Qualitative abstraction does not allow model to be ruled out

x_a

0 ┄┄┄┄┄┄┄┄

time

x_b

0 ┄┄┄┄┄┄┄┄

time

- $\dot{x}_a > 0$ • $\dot{x}_a < 0$
- $\dot{x}_b > 0$ • $\dot{x}_b < 0$

max_b

κ_b/γ_b

θ_b

0

θ_{a1} θ_{a2} κ_a/γ_a max_a

M^{11} M^{12} M^{13} M^{14} M^{15}

M^6 M^7 M^8 M^9 M^{10}

M^1 M^2 M^3 M^4 M^5

- $\dot{x}_a < 0$
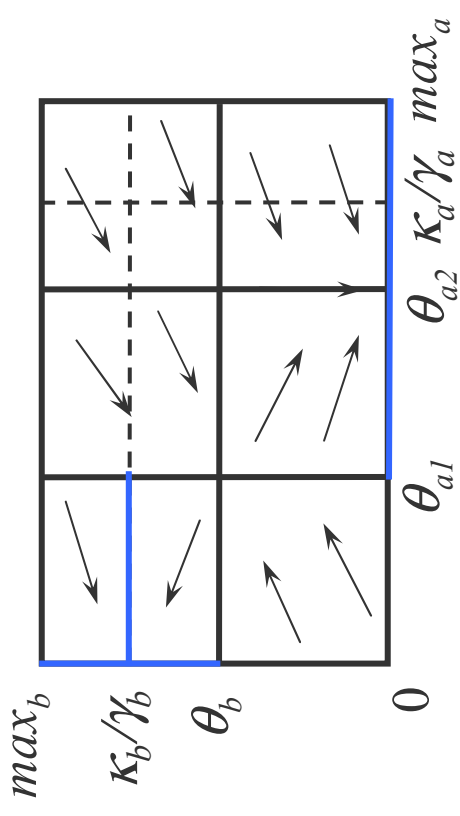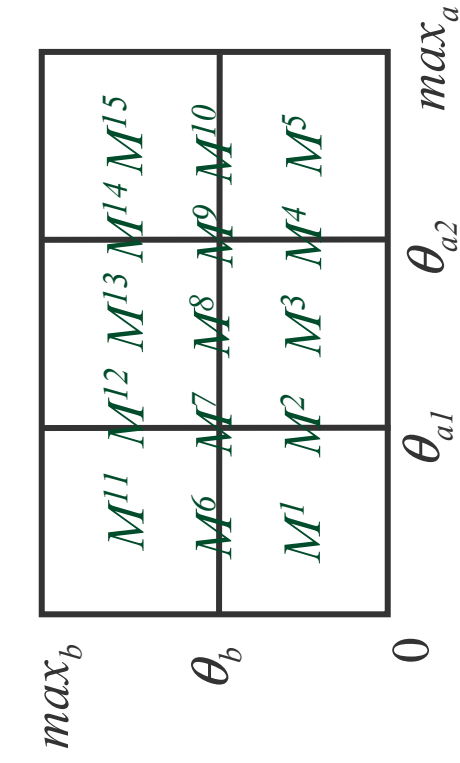- $\dot{x}_b$?

- $\dot{x}_a > 0$
- $\dot{x}_b > 0$

❖ Partition of phase space too coarse: need for finer-grained partition with unique derivative sign pattern in each domain

Refinement of phase-space partition

❖ Refinement of partition of phase space: from **mode domains** to **flow domains**
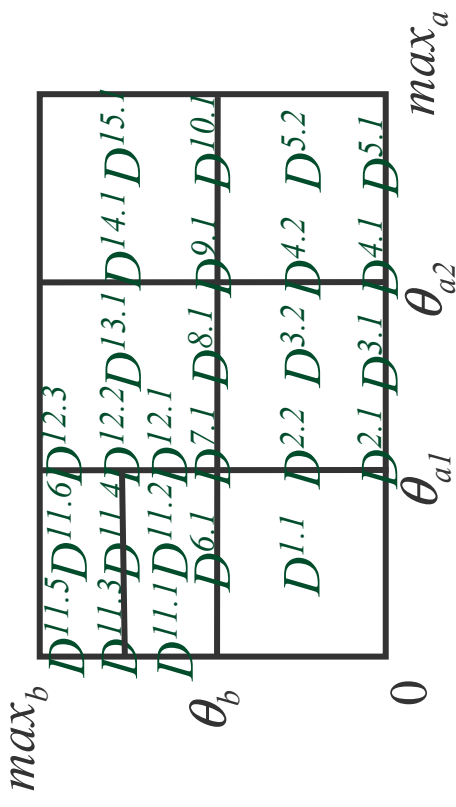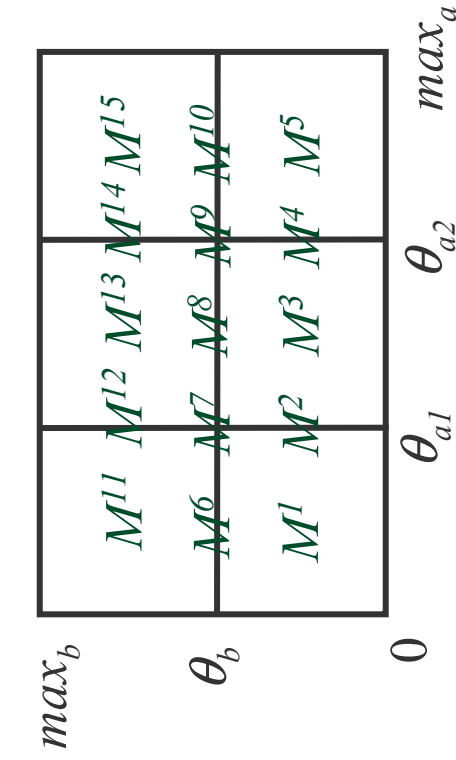
Repartitioning of mode domains by means of nullcline planes

Refinement of phase-space partition

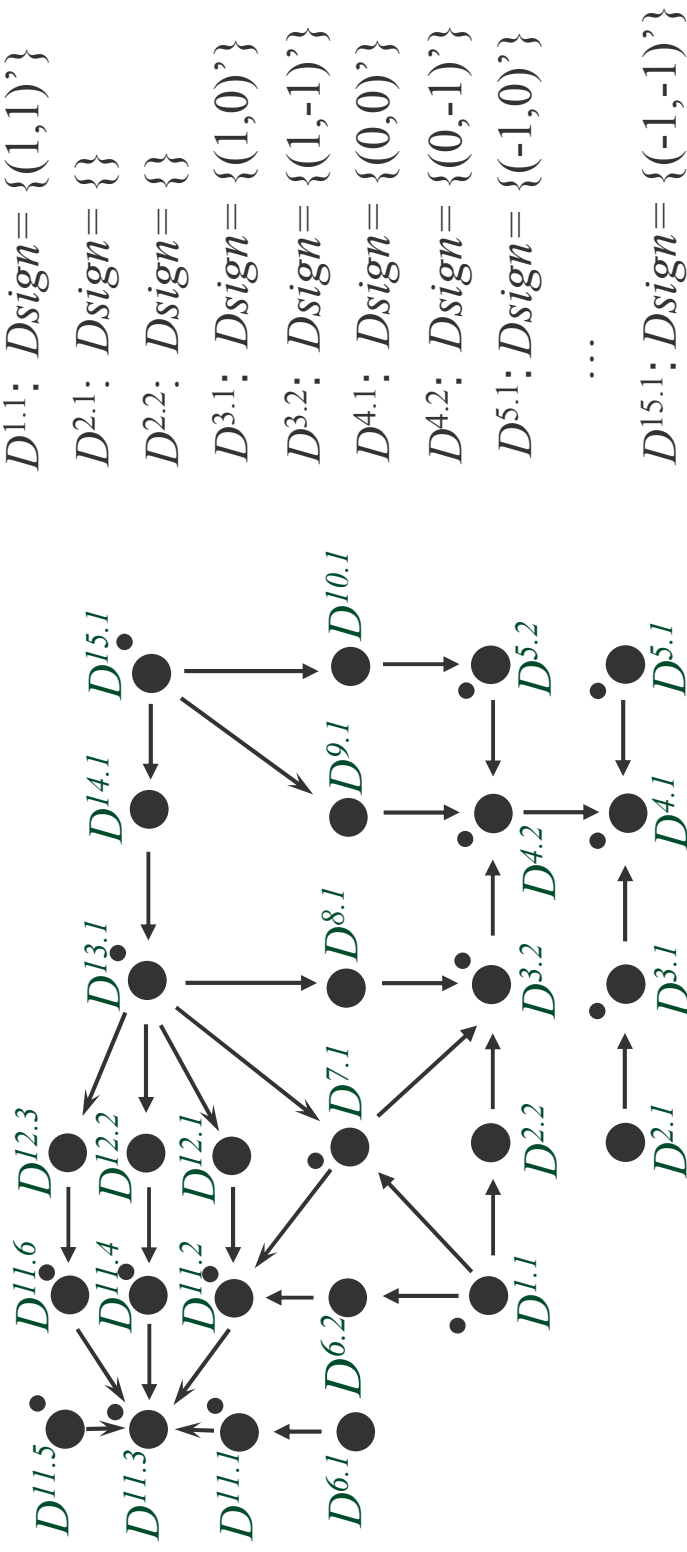❖ Refinement of partition of phase space: from **mode domains** to **flow domains**

Repartitioning of mode domains by means of nullcline planes

The mode domain diagram (left) with regions M^1 through M^{15}:

M^{11} M^{12} M^{13} M^{14} M^{15}

M^6 M^7 M^8 M^9 M^{10}

M^1 M^2 M^3 M^4 M^5

axes: max_b, θ_b, 0 (vertical); 0, θ_{a1}, θ_{a2}, max_a (horizontal)

The flow domain diagram (right):

$D^{11.5}$ $D^{11.6}$ $D^{12.3}$ $D^{14.1}$ $D^{15.1}$
$D^{11.3}$ $D^{11.4}$ $D^{12.2}$ $D^{13.1}$
$D^{11.1}$ $D^{11.2}$ $D^{12.1}$
$D^{6.1}$ $D^{7.1}$ $D^{8.1}$ $D^{9.1}$ $D^{10.1}$
$D^{1.1}$ $D^{2.2}$ $D^{3.2}$ $D^{4.2}$ $D^{5.2}$
$D^{2.1}$ $D^{3.1}$ $D^{4.1}$ $D^{5.1}$

axes: max_b, θ_b, 0 (vertical); 0, θ_{a1}, θ_{a2}, max_a (horizontal)
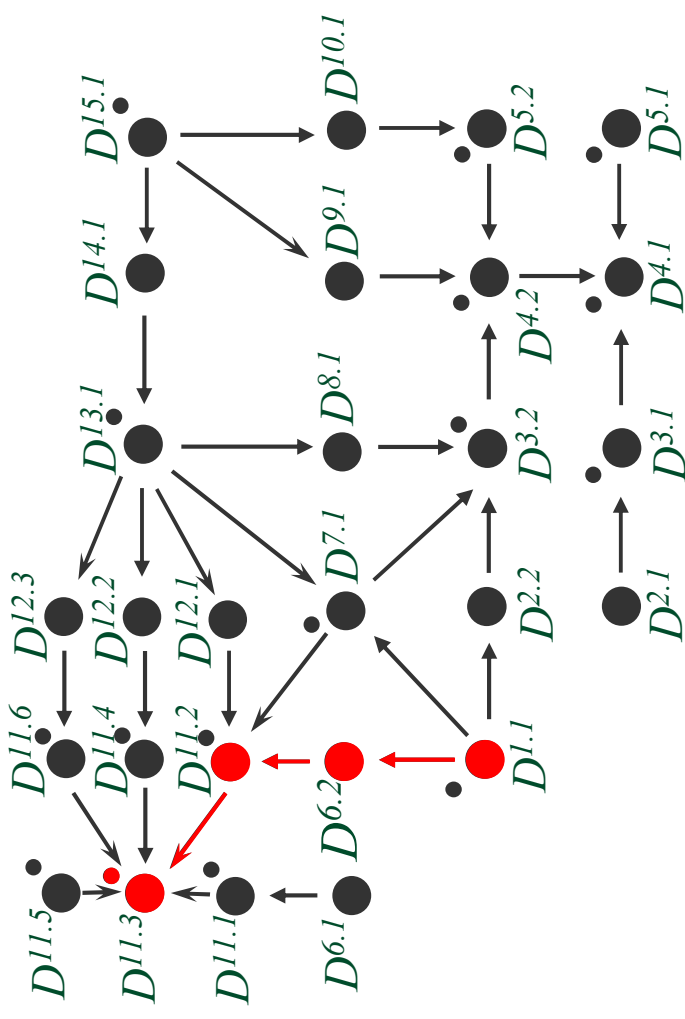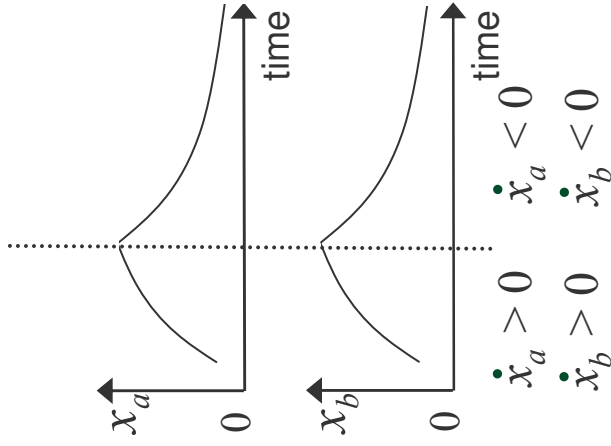
❖ Unique derivative sign pattern in every flow domain

$D^{1.1}$: $Dsign = \{(1,1)'\}$; ...; $D^{11.2}$: $Dsign = \{(-1,1)'\}$; $D^{11.3}$: $Dsign = \{(0,0)'\}$

Improvement of model validation

$D^{1.1}$: $Dsign = \{(1,1)'\}$

$D^{2.1}$: $Dsign = \{\}$

$D^{2.2}$: $Dsign = \{\}$

$D^{3.1}$: $Dsign = \{(1,0)'\}$

$D^{3.2}$: $Dsign = \{(1,-1)'\}$

$D^{4.1}$: $Dsign = \{(0,0)'\}$

$D^{4.2}$: $Dsign = \{(0,-1)'\}$

$D^{5.1}$: $Dsign = \{(-1,0)'\}$

$\ldots$

$D^{15.1}$: $Dsign = \{(-1,-1)'\}$

Graph nodes: $D^{15.1}$, $D^{14.1}$, $D^{13.1}$, $D^{12.3}$, $D^{12.2}$, $D^{12.1}$, $D^{11.6}$, $D^{11.4}$, $D^{11.2}$, $D^{11.5}$, $D^{11.3}$, $D^{11.1}$, $D^{6.1}$, $D^{6.2}$, $D^{10.1}$, $D^{9.1}$, $D^{8.1}$, $D^{7.1}$, $D^{5.2}$, $D^{4.2}$, $D^{4.1}$, $D^{3.2}$, $D^{3.1}$, $D^{2.2}$, $D^{2.1}$, $D^{5.1}$, $D^{1.1}$

UNIVERSITÉ JOSEPH FOURIER

INRIA RHÔNE-ALPES

Improvement of model validation

$D^{1.1}$: $Dsign=\{(1,1)^?\}$; $D^{6.2}$: $Dsign=\{\}$

$D^{11.2}$: $Dsign=\{(-1,1)^?\}$; $D^{11.3}$: $Dsign=\{(0,0)^?\}$

❖ Using fined-grained partition, the model can be safely ruled out

Model checking approach

❖ **Model checking** is automated technique for verifying that discrete transition system satisfies certain temporal properties

❖ CTL model checking framework:

- set of **atomic propositions** AP

- discrete transition system is **Kripke structure** $KS = \langle S, R, L \rangle$, where S set of states, R transition relation, L labeling function over AP

- temporal properties expressed in **Computation Tree Logic (CTL)**

  $p, \neg f_1, f_1 \wedge f_2, f_1 \vee f_2, f_1 \rightarrow f_2, EXf_1, AXf_1, EFf_1, AFf_1, EGf_1, AGf_1, Ef_1Uf_2, Af_1Uf_2,$ where $p \in AP$ and f_1, f_2 CTL formulas

❖ Computer tools are available to perform efficient and reliable model checking (*e.g.*, NuSMV, SPIN, CADP)
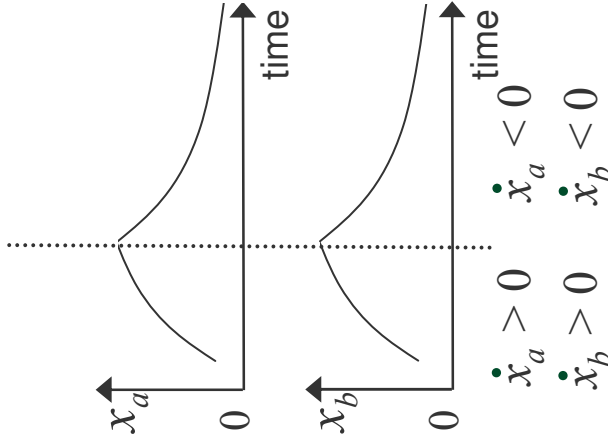
Validation using model checking

❖ Atomic propositions

$$AP = \{value_a = 0,\ value_a < \theta_a^1,\ ...,\ Dsign_a = \{-1\},\ Dsign_a = \{0\},...\}$$

❖ Observed property expressed in CTL

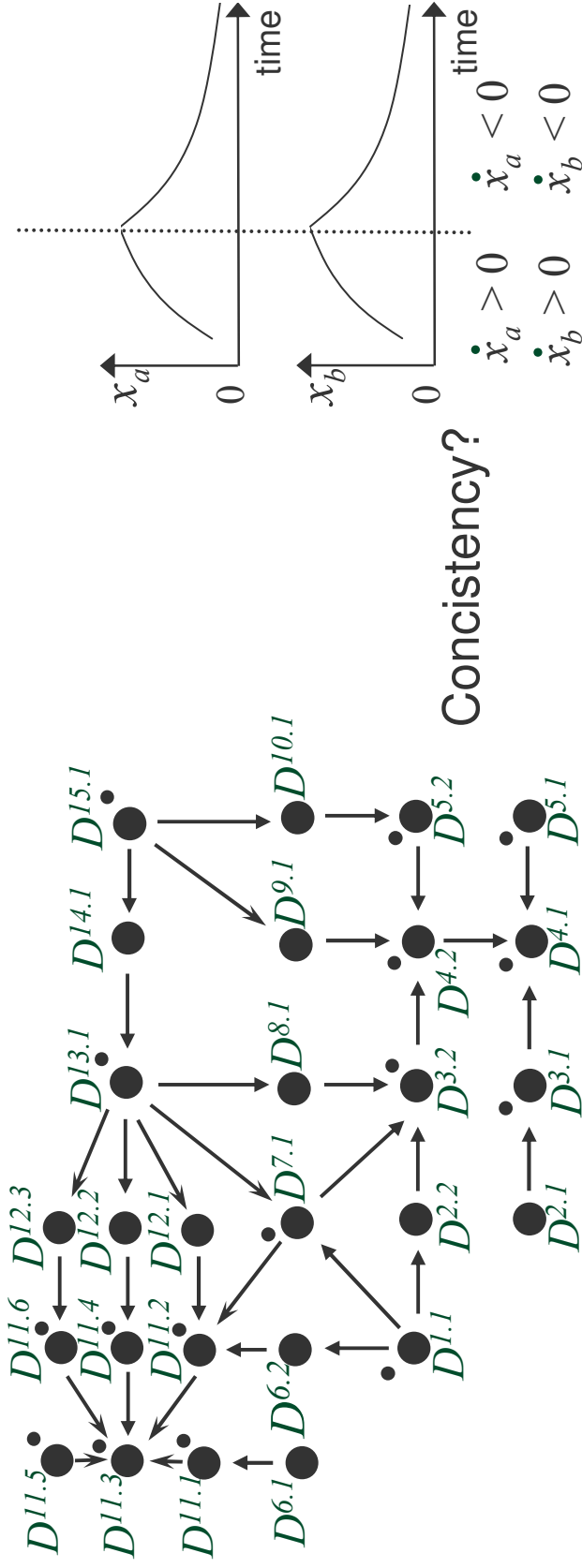There *Exists* a *Future* state where $Dsign_a=\{1\}$ and $Dsign_b=\{1\}$ and starting from that state,

there *Exists* a *Future* state where $Dsign_a=\{-1\}$ and $Dsign_b=\{-1\}$

EF$(Dsign_a=\{1\}\ \Lambda\ Dsign_b=\{1\}\ \Lambda$
EF$(Dsign_a=\{-1\}\ \Lambda\ Dsign_b=\{-1\}))$

Validation using model checking

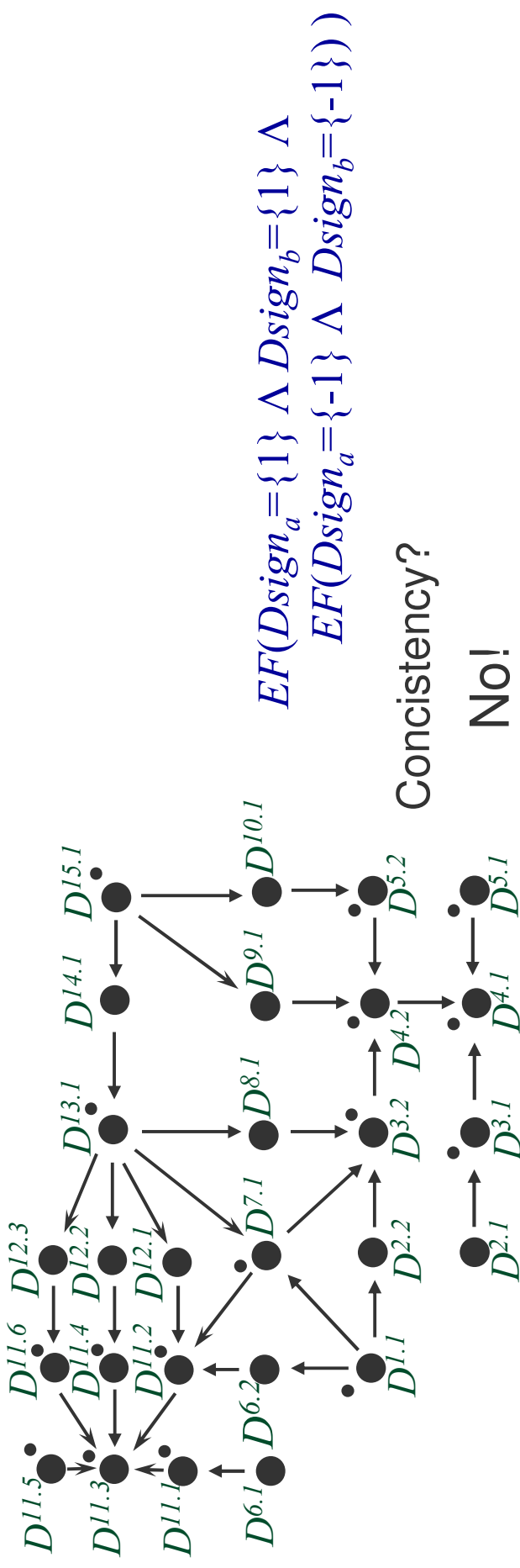❖ Check consistency between predictions and experimental data

Concistency?

$x_a > 0$ $x_a < 0$
$x_b > 0$ $x_b < 0$

$D^{1.1}$: $Dsign = \{(1,1)'\}$; $D^{6.1}$: $Dsign = \{\}$

$D^{11.2}$: $Dsign = \{(-1,1)'\}$; $D^{11.3}$: $Dsign = \{(0,0)'\}$

Validation using model checking

❖ Check consistency between predictions and experimental data

$EF(Dsign_a=\{1\} \wedge Dsign_b=\{1\} \wedge$
$EF(Dsign_a=\{-1\} \wedge Dsign_b=\{-1\}))$

Concistency?

No!

$D^{1.1}:$ $Dsign=\{(1,1)\}$; $D^{6.1}:$ $Dsign=\{\}$
$D^{11.2}:$ $Dsign=\{(-1,1)\}$; $D^{11.3}:$ $Dsign=\{(0,0)\}$

❖ Model ruled out by model checker

$D^{11.5}$ $D^{11.6}$ $D^{12.3}$
$D^{11.3}$ $D^{11.4}$ $D^{12.2}$
$D^{11.1}$ $D^{11.2}$ $D^{12.1}$
$D^{6.1}$ $D^{6.2}$ $D^{7.1}$ $D^{8.1}$ $D^{13.1}$ $D^{14.1}$ $D^{15.1}$
$D^{1.1}$ $D^{2.2}$ $D^{3.2}$ $D^{4.2}$ $D^{9.1}$ $D^{10.1}$
$D^{2.1}$ $D^{3.1}$ $D^{4.1}$ $D^{5.2}$
$D^{5.1}$

18

UNIVERSITE JOSEPH FOURIER
SCIENCES TECHNOLOGIE MEDECINE

INRIA
RHÔNE-ALPES

Genetic Network Analyzer (GNA)

❖ Tailored algorithms for symbolic computation of transition system implemented in new version of GNA
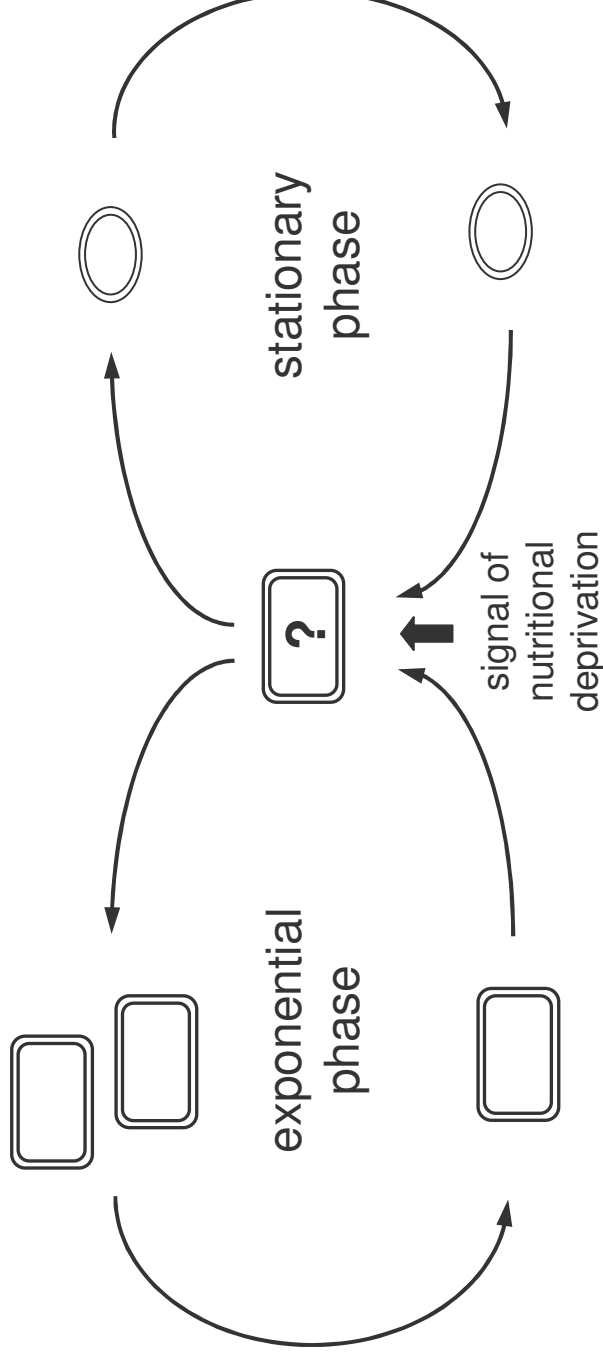
❖ Export functionalities to model checkers (NuSMV, CADP)
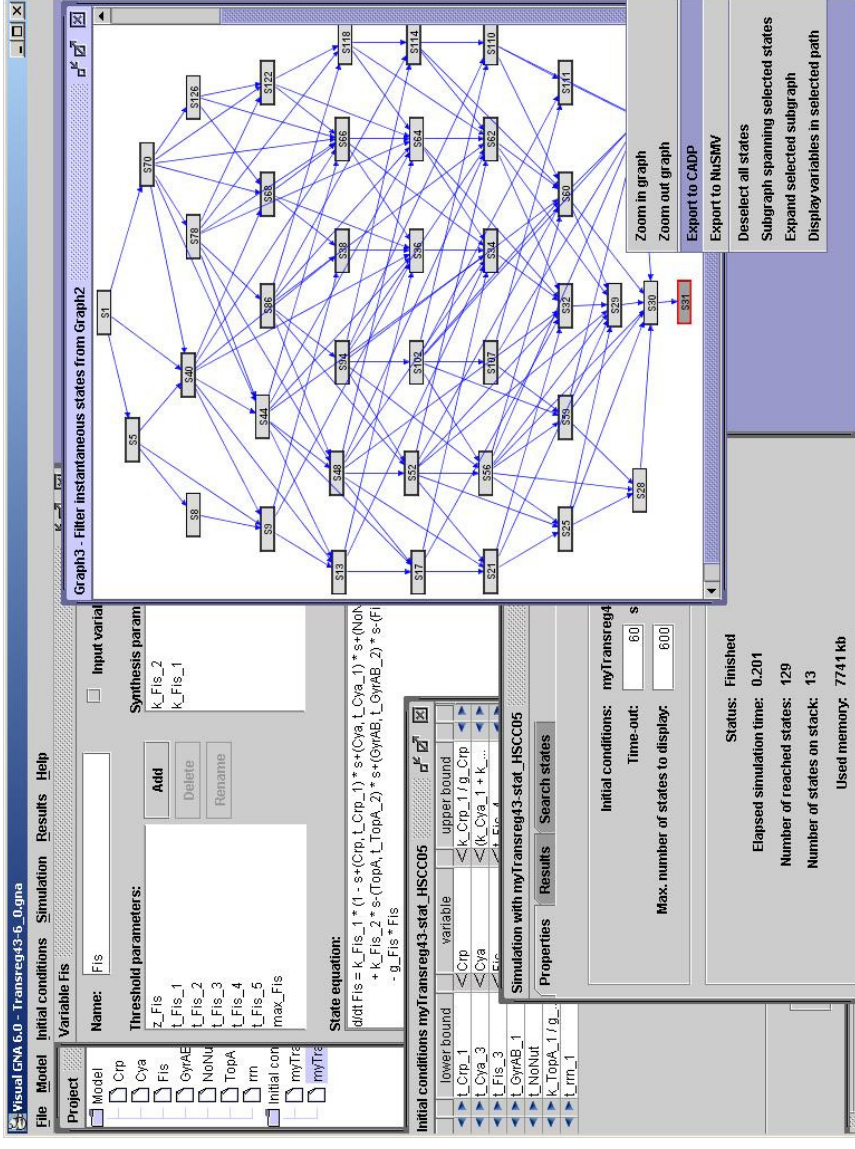
de Jong *et al.*, *Bioinformatics*, 03   Batt *et al.*, *Bioinformatics*, 05
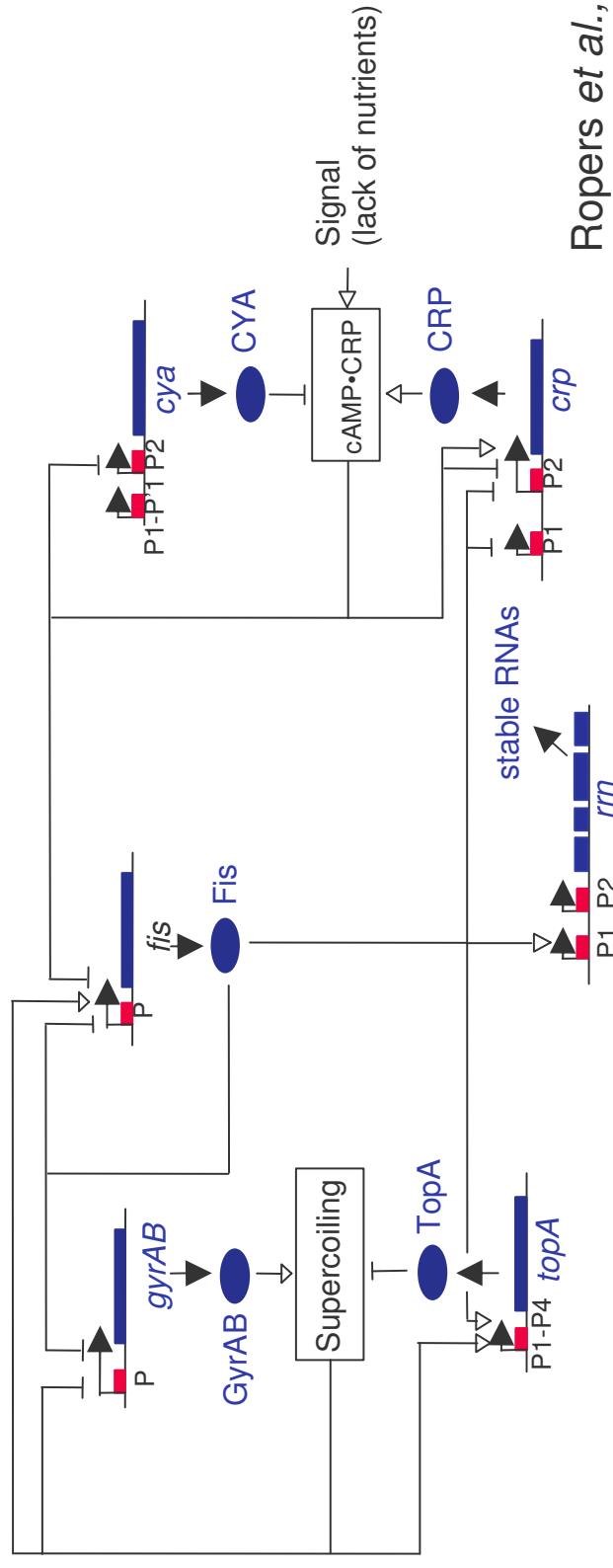
Nutritional stress response in *E. coli*

❖ In case of nutritional stress, *E. coli* population abandons growth and enters stationary phase

exponential phase

stationary phase

signal of nutritional deprivation

?

Nutritional stress response in *E. coli*

❖ In case of nutritional stress, *E. coli* population abandons growth and enters stationary phase


Ropers *et al.*, 05

❖ Decision to abandon or continue growth is controlled by complex genetic regulatory network

❖ Model: 7 PLDEs, 40 parameters and 54 inequality constraints

Validation of stress response model

❖ Qualitative simulation of nutrient starvation:

- 66 reachable domains (< 1s.)

- single attractor domain (asymptotically stable equilibrium point)

❖ Experimental data on Fis:
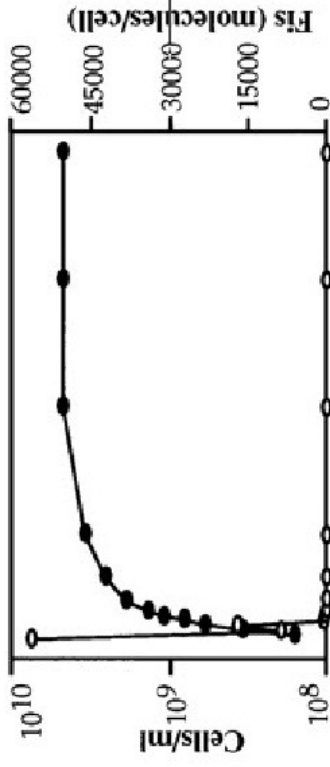
Fis concentration decreases and becomes steady in stationary phase

Ali Azam et al., J. Bacteriol., 99

❖ CTL formulation: $EF(Dsign_{fis} = \{-1\} \land EF(Dsign_{fis} = \{0\} \land value_{rrn} < \theta_{rrn}))$

❖ Model checking with NuSMV: property true (< 1s.)

Validation of stress response model

❖ Other properties:

- DNA supercoiling decreases during transition to stationary phase
  Balke and Gralla, *J. Bacteriol.*, 87

  $EF((dsign_{gyrAB} = \{-1\} \lor dsign_{topA} = \{1\}) \land value_{rrn} < \theta_{rrn})$ False

- *rrn* transcription increases independently of Fis following upshift
  Appleman et al., *J. Bacteriol.*, 99

  $E(value_{fis} < \theta^3_{fis} \ U \ Dsign_{rrn} = \{1\})$ False

- *cya* transcription is negatively regulated by the complex cAMP-CRP
  Kawamukai et al., *J. Bacteriol.*, 85

  $AG(value_{crp} > \theta^3_{crp} \land value_{cya} > \theta_{cya} \land value_s > \theta_s$
  $\rightarrow EF \ Dsign_{cya} = \{-1\})$ True

❖ Need for **model-driven experiments**:

time-series measurement of concentrations in parallel at high sampling rate

23

Conclusions

❖ Refinement of qualitative simulation method to yield predictions **suitable for comparison** with experimental data

❖ Combination with model-checking techniques to achieve **automatic** and **efficient** comparison

❖ Approach **implemented** and **applied** to nutritional stress response in *E. coli*

❖ Model-checking used in combination with different formalisms

- generalized logical models Bernot *et al., J. Theor. Biol.,* 04
- concurrent systems Chabrier *et al., Theor. Comput. Sci.,* 04, Eker *et al., PSB02*
- hybrid automata Ghosh *et al., HSCC03*

❖ Further work: integration of tailored model checker in GNA

❖ Thanks for your attention!