Identification of deterministic piecewise affine models of genetic regulatory networks

Giancarlo Ferrari-Trecate

Dipartimento di Informatica e Sistemistica (DIS),
Università degli Studi di Pavia,
Italy

giancarlo.ferrari@unipv.it

Thanks to: O. Bernard, D. Chieppi, S. Druhle, H. de Jong, R. Porreca

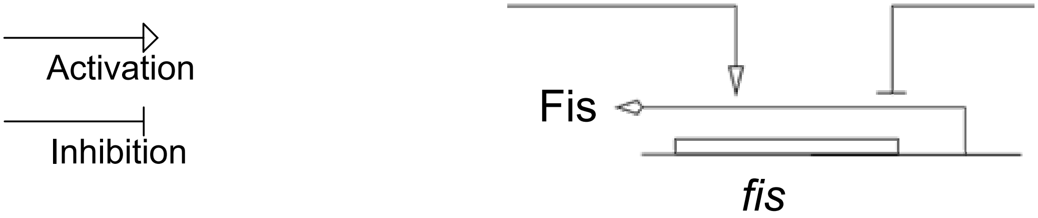
Overview

1. Basics on Genetic Regulatory Networks (GRNs) and their identification
2. PieceWise Affine (PWA) models of GRNs
3. Data-based reconstruction of GRNs
   - Pitfalls of general methods for PWA system identification
   - Towards gray-box identification of GRNs
     - Switch detection
     - Threshold reconstruction
4. A case study: identification of *E. coli* carbon starvation response
5. Conclusions

Genetic regulatory networks

- **GRNs** underlie functioning and development of living organisms
  - *Components:* genes, proteins, small molecules, and their mutual regulatory interactions

Genes



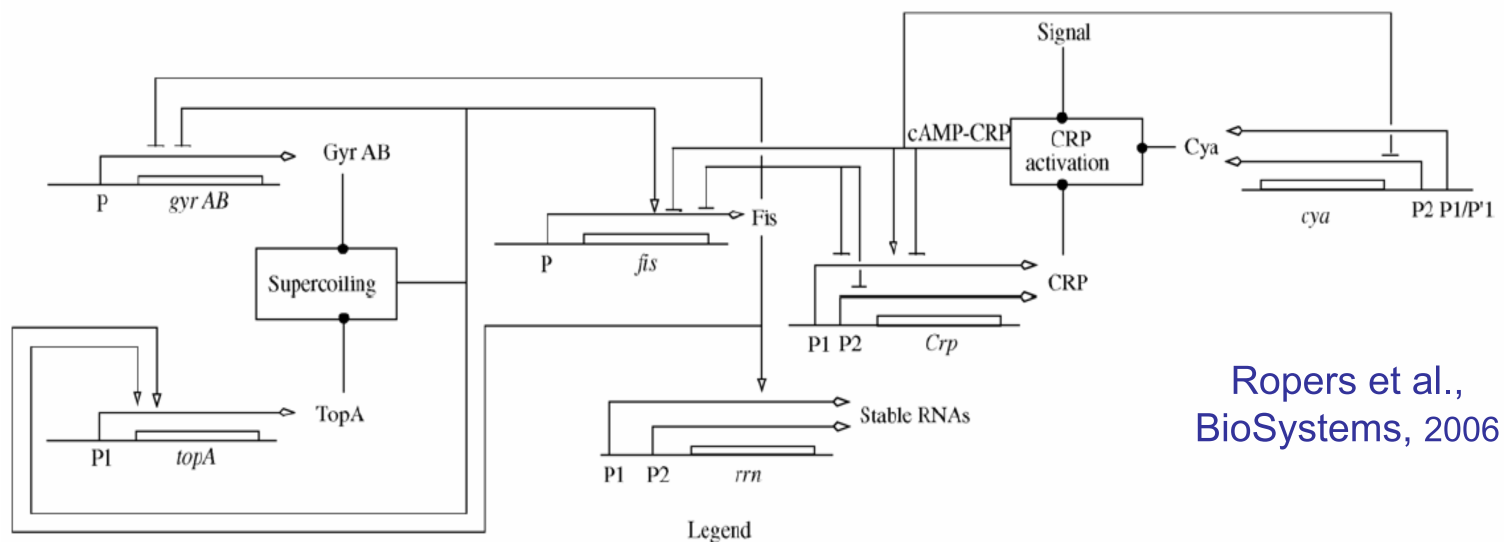
- Gene: dynamical system coding for a molecule (e.g. a protein)
- Genes are regulated by the concentration of proteins present in the cell
  - Genes can be turned on and off

Genetic regulatory networks

- **GRNs** underlie functioning and development of living organisms
  - *Components:* genes, proteins, small molecules and their mutual regulatory interactions

- GRNs are usually **large** (many genes) and **complex** (feedback loops)

GRN governing *E. coli* carbon starvation response

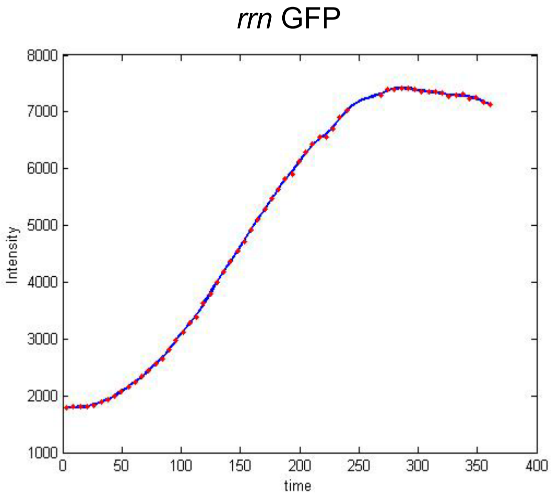


Ropers et al., BioSystems, 2006

Gene expression data

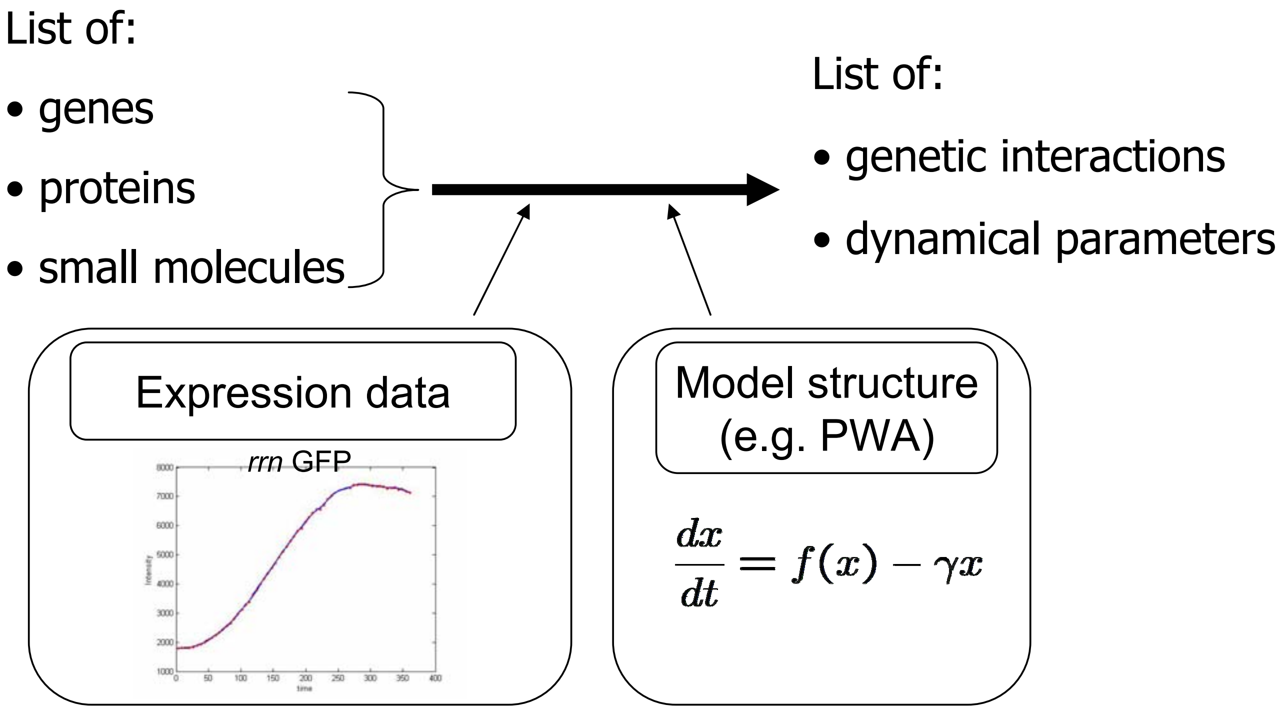
- Experimental techniques in biology have led to the production of enormous amount of data on the dynamics of gene expression:
  - DNA microarrays
  - gene reporter systems

Time-series measurement of fluorescence or luminescence



Data-driven modeling of GRNs

- System identification problem: derive a model of the regulatory interactions according to measurements and model structure



List of:

- genes
- proteins
- small molecules

List of:

- genetic interactions
- dynamical parameters

Expression data

rrn GFP

Model structure (e.g. PWA)

$$\frac{dx}{dt} = f(x) - \gamma x$$

Gene reporter systems ⇒ adequate sampling time to capture GRN dynamics

State of the art

Classes of dynamical models that were used for modeling genes and GRNs:

- Linear (Gardner et al., Science 301, 2003)
  → only valid near an equilibrium point
- Nonlinear smooth (Jaeger et al., Nature 430, 2004)
  → more adequate description but difficult to use for identification
- PieceWise Affine (PWA)
  → compromise between linear and non-linear
  - Introduced by Glass and Kauffman in the 1970s
  - de Jong et al., Bull. Math. Biol. 66, 2004
  - Ghosh and Tomlin, Syst. Biol. 1, 2004
  - Batt et al., HSCC05, Vol. 3414 of LNCS, 2005

  - → tools for analysis and abstractions available
  - → identification methods for PWA systems available

PWA models of GRNs

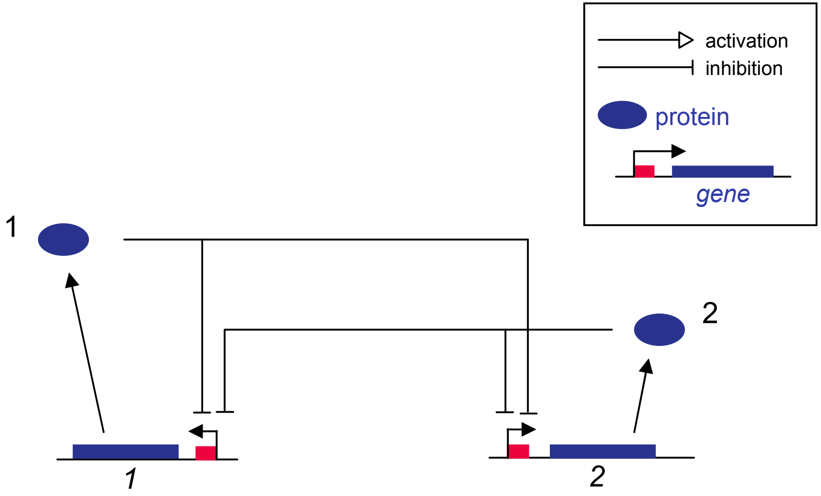
Consider a GRN composed by n genes

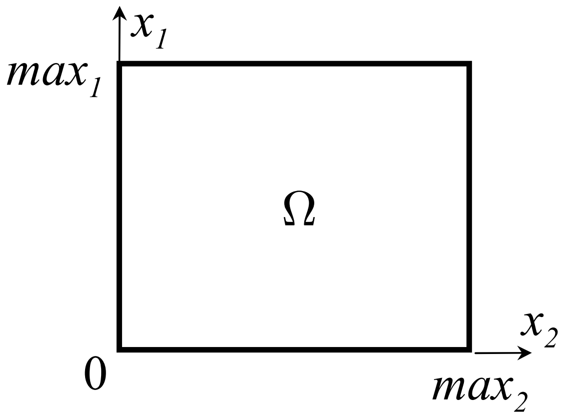
- State vector: $x = \underline{[x_1, x_2, \ldots, x_n]} \in \Omega$

gene product concentrations

- State set $\Omega \subset \mathbb{R}^n_{\geq 0}$: hyperrectangle including the origin

Toy example





PWA models of GRNs

synthesis rate ≥ 0

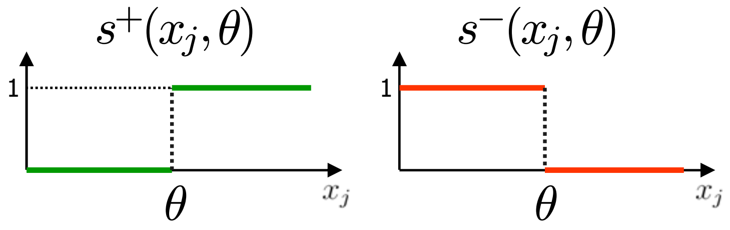
GRN dynamics: $\dot{x}_i = f_i(x) - g_i(x)x, \quad i = 1, \dots, n$

degradation rate > 0

$$f_i(x) = \sum_{l \in L_i} \alpha_{il} b_{il}(x)$$

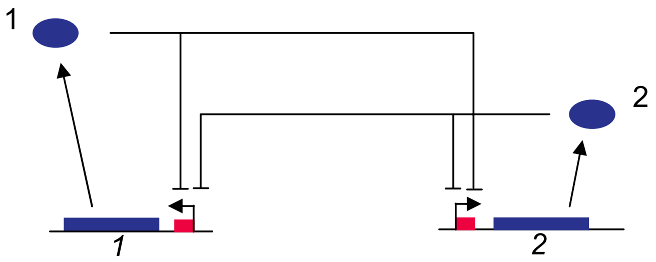
$$g_i(x) = \sum_{l \in \tilde{L}_i} \tilde{\alpha}_{il} \tilde{b}_{il}(x)$$

0/1-valued polynomials of step functions



θ : switching threshold

Toy example



$$\dot{x}_1 = \alpha_{11} b_{11}(x) - \tilde{\alpha}_{11} x_1$$

$$\dot{x}_2 = \alpha_{21} b_{21}(x) - \tilde{\alpha}_{21} x_2$$

$$b_{11} = s^-(x_1, \theta_1^1) s^-(x_2, \theta_2^1)$$

$$b_{21} = s^-(x_1, \theta_1^2) s^-(x_2, \theta_2^2)$$

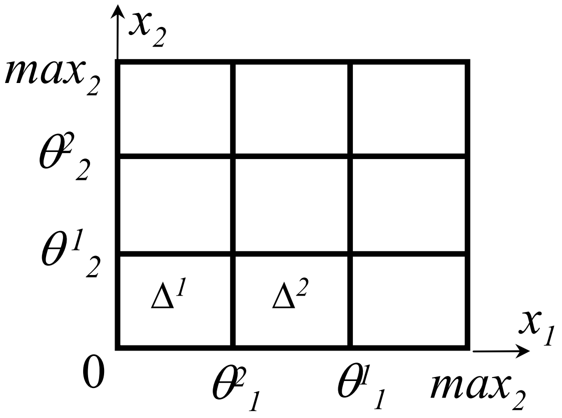
PWA models of GRNs

- All thresholds split Ω into hyperrectangular domains $\{\Delta^j\}_{j=1}^s$
- Step functions are constant on each domain $\Rightarrow$ PWA system

$$\dot{x} = \mu^j - \nu^j x \quad \text{if} \quad \lambda(x) = j$$

- $\mu^j = \begin{bmatrix} \mu_1^j & \dots & \mu_n^j \end{bmatrix}^T \geq 0,\ \nu^j = \mathrm{diag}(\nu_1^j, \dots, \nu_n^j) > 0$
- $\lambda(x) = j \Leftrightarrow x \in \Delta^j$: switching function

Toy example



$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{cases} \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} - \begin{bmatrix} \tilde{\alpha}_{11} & 0 \\ 0 & \tilde{\alpha}_{21} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, & \text{if } \lambda(x) = 1 \\ \begin{bmatrix} \alpha_{11} \\ 0 \end{bmatrix} - \begin{bmatrix} \tilde{\alpha}_{11} & 0 \\ 0 & \tilde{\alpha}_{21} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, & \text{if } \lambda(x) = 2 \\ \vdots \end{cases}$$

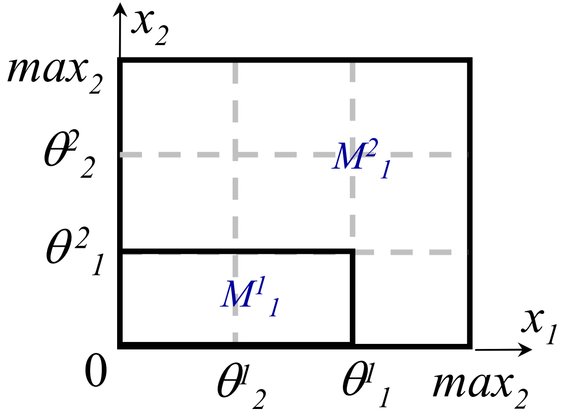
PWA model of a molecule concentration

Dynamics of the *i*-th molecule concentration:

$$\dot{x}_i = \kappa_i^j - \gamma_i^j x_i \quad \text{if} \quad x \in M_i^j$$

- $\{M_i^j\}_{j=1}^{s_i}$: molecule domains (regions in Ω where the *i*-th concentration obeys to the same dynamics)
- Inputs: $x_p, \ p \neq i$

Toy example



Dynamics of x_1:

$$\dot{x}_1 = \begin{cases} \alpha_{11} - \tilde{\alpha}_{11} x_1, & \text{if } x \in M_1^1 \\ -\tilde{\alpha}_{11} x_1, & \text{if } x \in M_1^2 \end{cases}$$

Standing assumption: no sliding-mode behaviors

Data model

Discrete-time model for the *i*-th molecule concentration:

$$x_i(k+1) = \tilde{\kappa}_i^j - \tilde{\gamma}_i^j x_i(k) + \eta_i(k) \quad \text{if} \quad x(k) \in M_i^j$$
$$y_i(k) = x_i(k) + \xi_i(k)$$

- $\tilde{\kappa}_i^j = (\kappa_i^j \backslash \gamma_i^j)(1 - e^{-\gamma_i^j T})$, $\tilde{\gamma}_i^j = -e^{\gamma_i^j T}$: rate parameters
- T: sampling time
- η_i, ξ_i : noise
- $y_i(k)$: measured data

Common data models:

- PieceWise Autoregressive eXogenous (PWARX): $\xi_i = 0$
- PWA Output-Error (PWA-OE): $\eta_i = 0$

Identification of GRNs

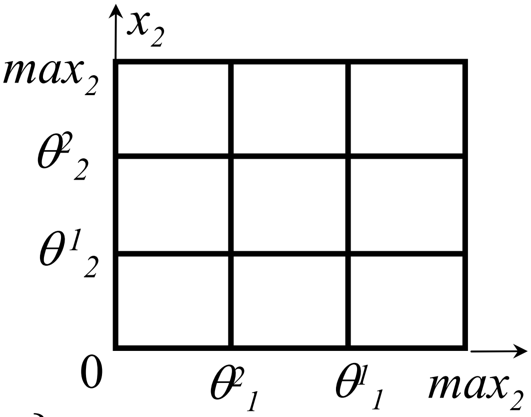
PWA discrete-time model of the GRN:

$$\begin{aligned} x_i(k+1) &= \tilde{\kappa}_i^j - \tilde{\gamma}_i^j x_i(k) + \eta_i(k) \quad \text{if} \quad x(k) \in M_i^j \\ y_i(k) &= x_i(k) + \xi_i(k) \\ i &= 1, \ldots, n \end{aligned}$$

Identification problem: reconstruct

- the number of modes
- all rate parameters
- all switching thresholds

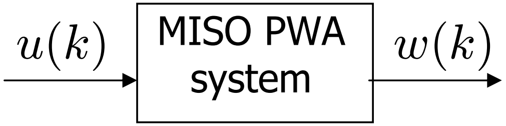
from the dataset $\{y_i(k), k = 1, \ldots, N, i = 1 : \ldots, n\}$



Can one use available algorithms for the identification of PWA models ?

Input-output PWA models

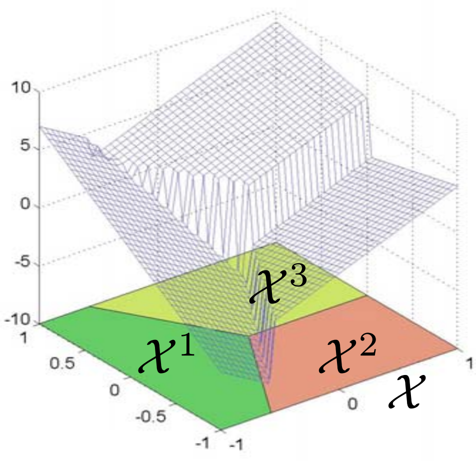
PWARX / PWA-OE models considered in hybrid identification:



$$\begin{aligned} z(k+1) &= \phi^j \begin{bmatrix} r(k)' & 1 \end{bmatrix}' + \eta(k) \quad \text{if} \quad r(k) \in \mathcal{X}^j \\ w(k) &= z(k) + \xi(k) \end{aligned}$$

- $r(k) = \begin{bmatrix} z(k) & \cdots & z(k-n_a) & u(k)' & \cdots & u(k-n_b)' \end{bmatrix}'$
- $\{\mathcal{X}^j\}_{j=1}^{\tilde{s}}$:polyhedral partition of the polytope $\mathcal{X}$

PWA models for a single molecule concentration fall within this class



Identification of I/O PWA models

PWARX / PWA-OE models considered in hybrid identification:

$$\begin{aligned} z(k+1) &= \phi^j \left[\begin{array}{cc} r(k)' & 1 \end{array} \right]' + \eta(k) \quad \text{if} \quad r(k) \in \mathcal{X}^j \\ w(k) &= z(k) + \xi(k) \end{aligned}$$

Data set = noisy samples

$\mathcal{N} = \{(r(k), w(k))\}_{k=1}^{N}$

- Common assumptions:

1. known model orders
2. known regressor set $\mathcal{X}$

- Estimate:

1. the number $\tilde{s}$ of modes
2. the parameters $\{\phi^j\}_{j=1}^{\tilde{s}}$
3. the regions $\{\mathcal{X}^j\}_{j=1}^{\tilde{s}}$

PWARX system identification:

(Bemporad et al., 2005), (Vidal et al., 2005), (Juloski et al., 2005), (Ferrari-Trecate et al., 2003), …

PWA-OE system identification:

(Juloski & Weiland, 2006), (Rosenqvist & Karlström, 2006)

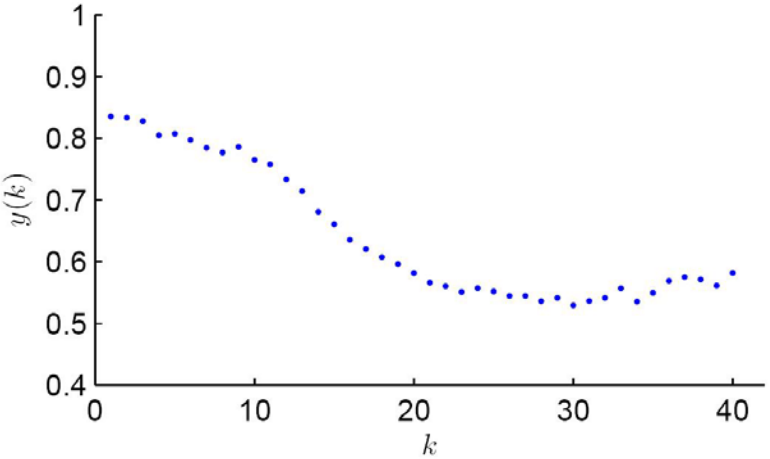
Software: Hybrid Identification Toolbox

HIT Hybrid Identification Toolbox

Pitfalls of available methods

Existing identification methods are generic in nature and do not exploit features of PWA models of GRNs

Example 1: Switch detection from noisy measurements



- Very challenging problem for general PWARX / PWA-OE models
- Much easier for PWA models of GRNs

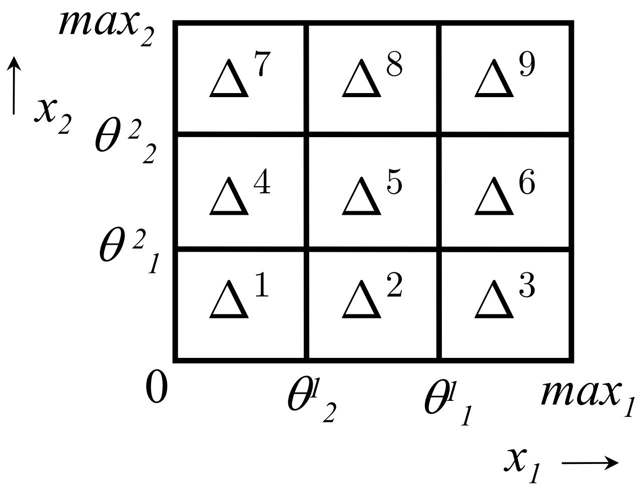
Pitfalls of available methods

Existing identification methods do not take into account constraints of PWA models of GRNs

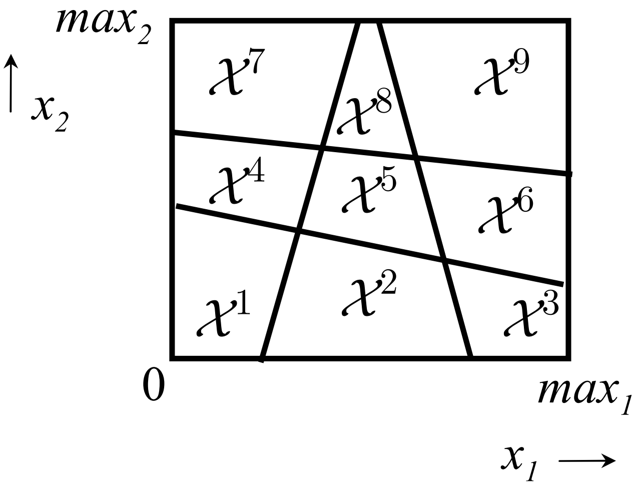
Example 2: switching thresholds ⇒ hyperrectangular domains

Neglecting this kind of information ...

True domains



Identified domains

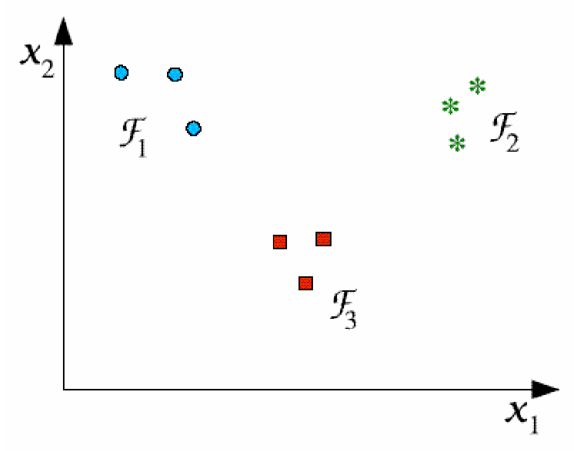


The concept of threshold associated to a concentration variable is lost

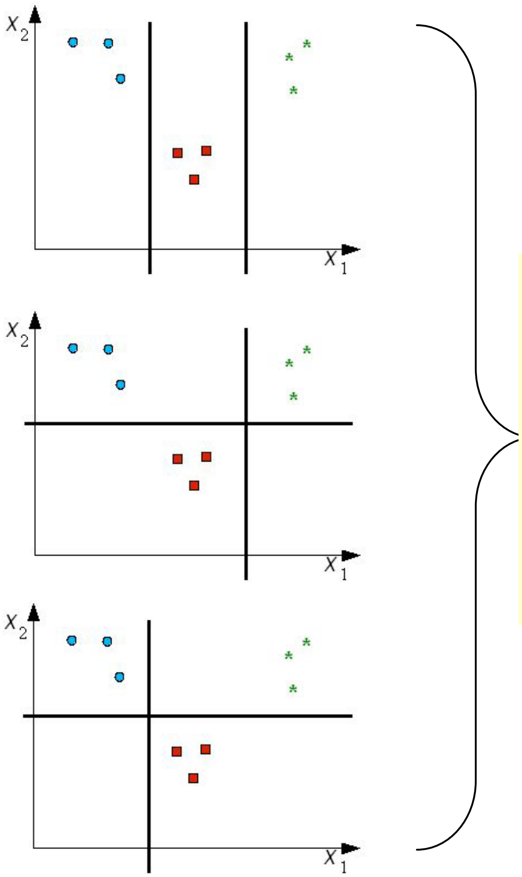
Pitfalls of available methods

Existing hybrid identification methods produce a single result but data are often scarce and multiple models might be plausible

Example: expression data in three domains



Problem: find thresholds separating domains



Three "minimal" combinations of thresholds

All of them should be produced !

Identification of PWA models of GRNs

Our approach: gray-box identification

1) Detection of switches in gene expression data

2) Estimation of the number of modes and attribution of the measurements to mode data sets

3) Reconstruction of
- thresholds on concentration variables
- all “minimal” combinations of thresholds consistent with the data

4) Estimation of kinetic parameters for all models generated in point 3

- Step 2 is currently under study
- Step 4 is easy (LS on each mode data set)

Next:
- two algorithms for step 1
- a procedure for step 3

Switching index

(Porreca et al., 2006)

PWA-OE model for the *i*-th molecule:

$$x(k+1) = \tilde{\kappa}^j - \tilde{\gamma}^j x(k) \quad \text{if} \quad x(k) \in M^j$$
$$y(k) = x(k) + \xi(k), \; \xi(k) \sim WGN(0, \sigma_n^2)$$

- $\{M^j\}_{j=1}^s$: molecule domains

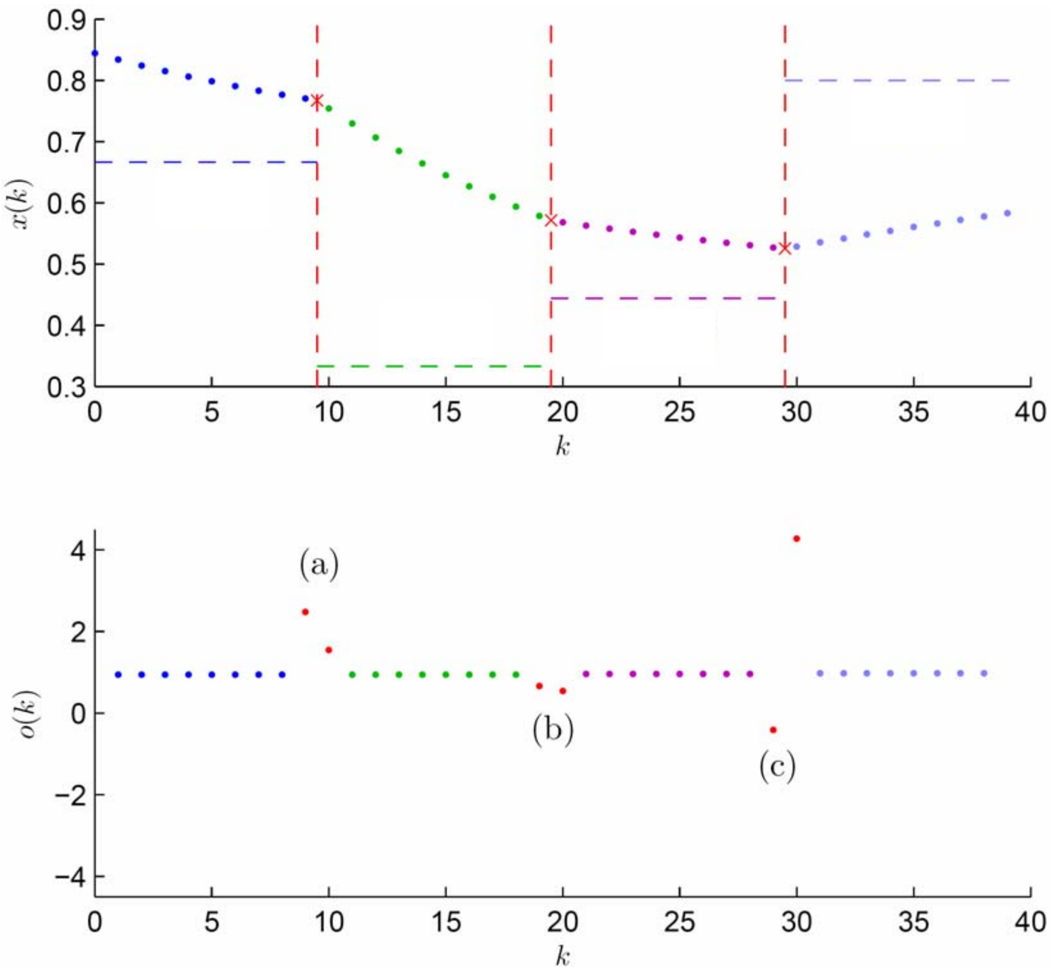
Switching index:

$$\boxed{o(k) = \frac{x(k+1) - x(k)}{x(k) - x(k-1)}}$$

The index emphasizes switches:

- if $x(k-1), \; x(k), \; x(k+1)$ belong to the same molecule domain for $k = k_a, \ldots, k_b$, then $o(k)$ is constant
- otherwise, it has a varying profile

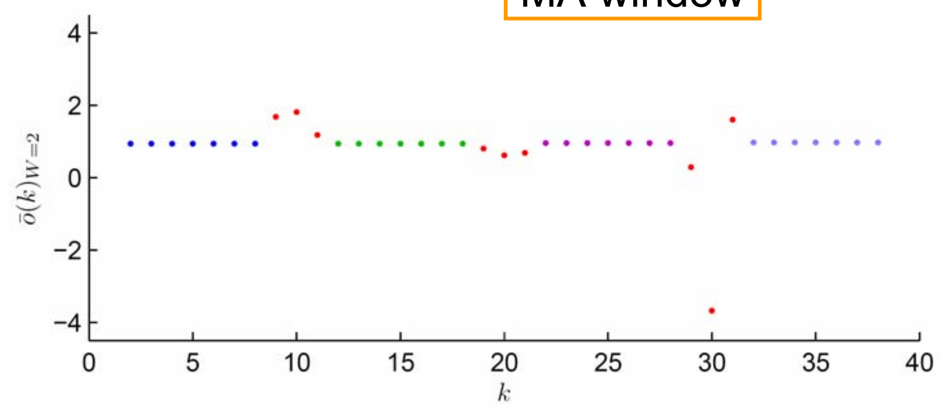
Behavior of the switching index

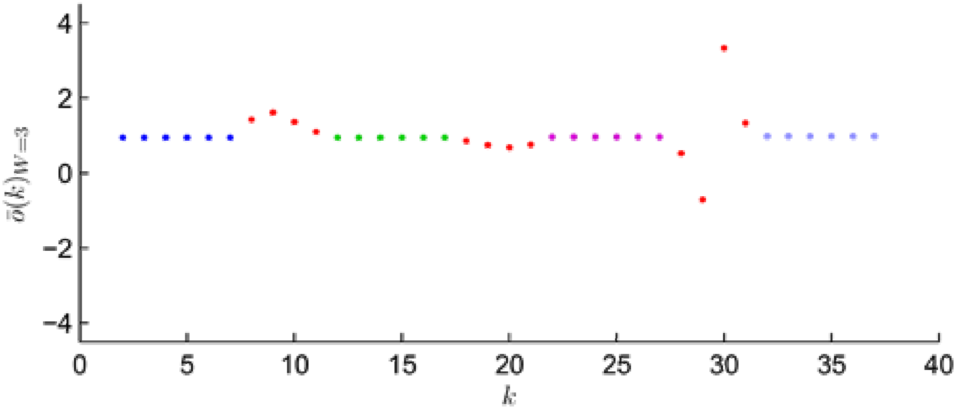


Moving Average (MA) switching indexes

$$\boxed{\bar{o}(k) = \frac{\bar{x}(k+1)-\bar{x}(k)}{\bar{x}(k)-\bar{x}(k-1)} = \frac{x(k+W)-x(k)}{x(k\underbrace{+W}_{\text{MA window}}-1)-x(k-1)}}$$

$$\bar{x}(k) = \frac{1}{W-2}\sum_{i=1}^{W-2} x(k+i)$$





Data-based indexes

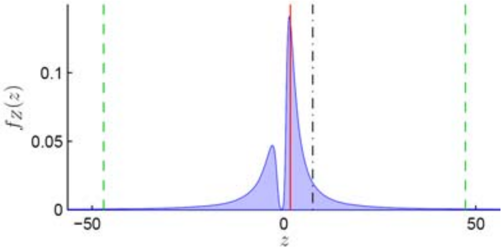
Data-based MA switching index: $\tilde{\bar{o}}(k) = \dfrac{y(k+W) - y(k)}{y(k+W-1) - y(k-1)}$

Ratio of two Gaussian random variables

$$Z = \frac{X_1}{X_2} \qquad \begin{array}{l} X_1 = y(k+W) - y(k), \quad X_1 \sim N\left(x(k+W) - x(k),\, 2\sigma_n^2\right) \\ X_2 = y(k+W-1) - y(k-1), \quad X_2 \sim N\left(x(k+W-1) - x(k-1),\, 2\sigma_n^2\right) \end{array}$$

Modified Cauchy distribution

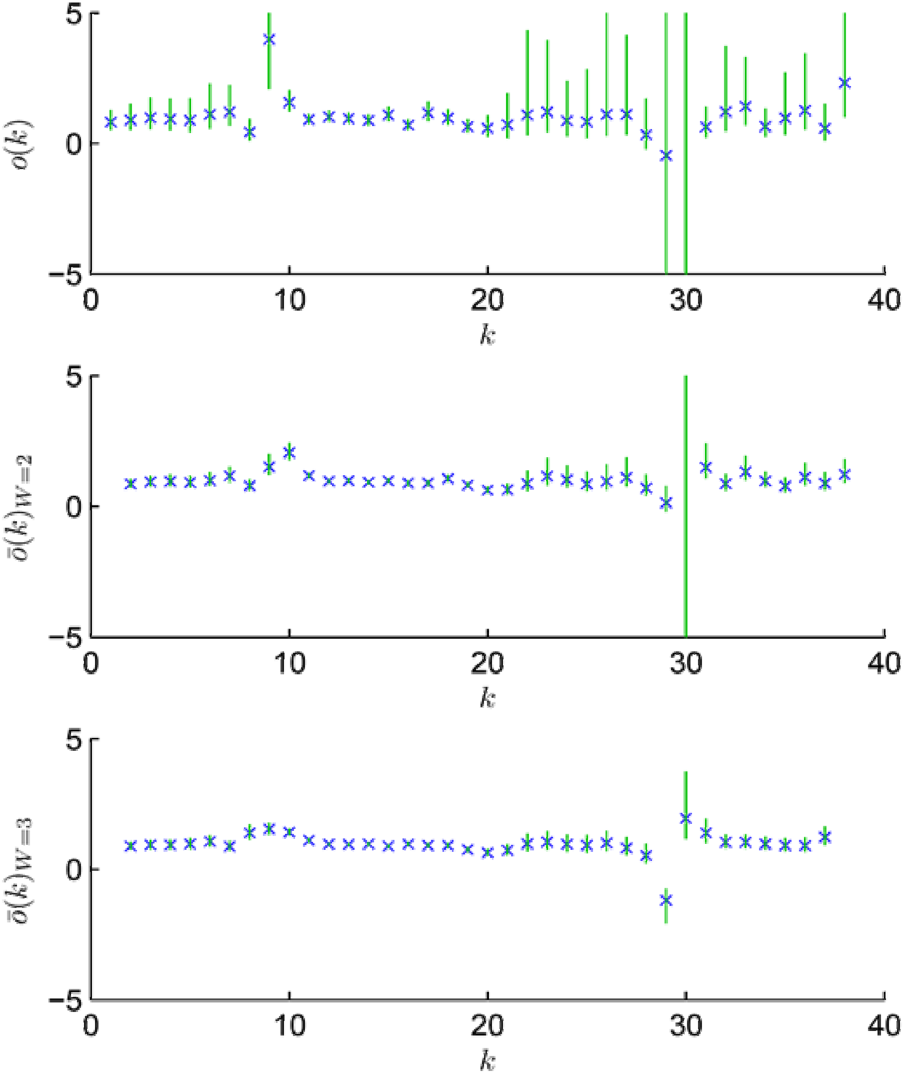
- undefined mean and variance



Fieller's theorem allows one to compute the α-level confidence sets for $\tilde{\bar{o}}(k)$ in closed form

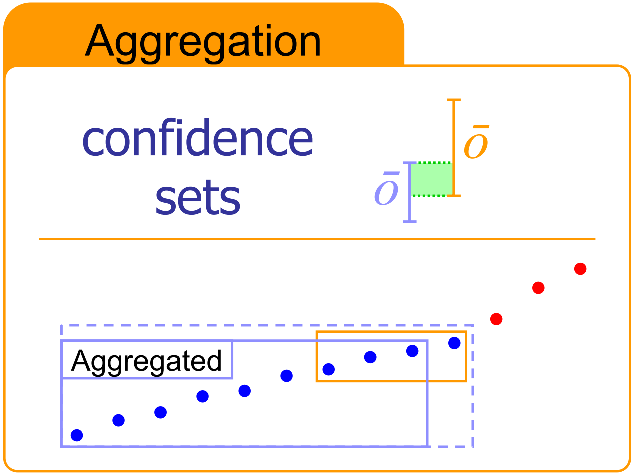
The higher W the smaller confidence sets

Data-based indexes



Switch detection algorithm

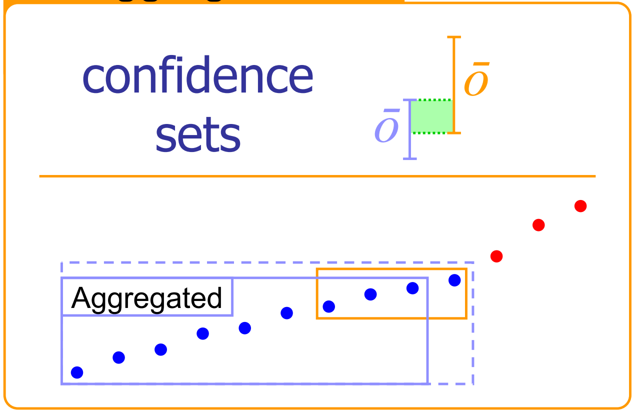
Key idea: aggregation rule based on confidence sets computed on different MA windows



Switch detection algorithm

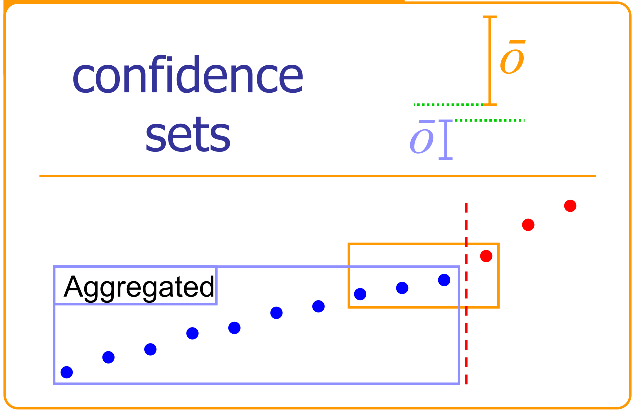
Key idea: aggregation rule based on confidence sets computed on different MA windows

Aggregation



confidence sets

Switch detection



confidence sets

Further features of the complete algorithm

(Porreca et al., 2006)

- re-inizialization after the detection of a switch
- backtracking for improving switch detection
- *ad hoc* handling of confidence sets of infinite length

Switch detection based on nonlinear estimation

Exponential model of the data (j-th mode):

$$y(k) = \frac{\kappa^j}{\gamma^j} - \left(\frac{\kappa^j}{\gamma^j} - x(k_0)\right) e^{-\gamma^j (k-k_0)T} + \xi(k)$$

Switch detection strategy:

- estimate $\hat{\kappa}^j, \ \hat{\gamma}^j, \ \hat{x}(k_0)$ using aggregated measures up to the time k_P
- hypothesis test:
    - H_0: $y(k_P+1)$ belongs to the same mode;
    - I_α: α-level confidence interval for $y(k_P+1)$ under H_0,
- switch detection rule: $\boxed{y(k_P+1) \notin I_\alpha}$

Comparison of the methods

Results based on extensive simumlations

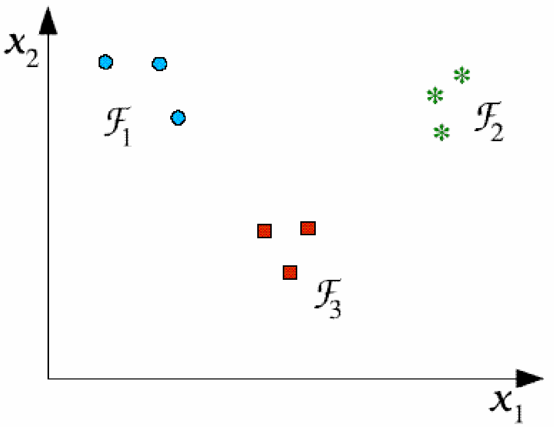
- Classification accuracy
- Molecule domain fragmentation

σ_n		switching indexes	nonlinear estimation
10^{-5}	accuracy	97,1%	75,3%
	fragmentation	4,4%	34,4%
10^{-4}	accuracy	93,8%	80,7%
	fragmentation	5,2%	26,9%
10^{-3}	accuracy	69,7%	69,7%
	fragmentation	16,4%	30,7%
10^{-2}	accuracy	22,3%	63,8%
	fragmentation	34,3%	15,2%

Reconstruction of switching thresholds

Assume that in early stages of identification:

- the number of modes has been estimated
- data have been attributed to modes of operation (i.e. data have been partitioned into mode data sets $\mathcal{F}_1, \ldots, \mathcal{F}_s$)



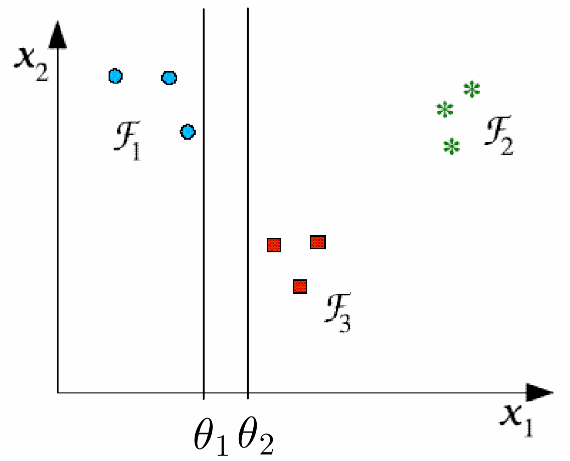
- Switching thresholds: axis-parallel (ap-) hyperplanes
- A set of switching thresholds consistent with the data must separate all pairs $(\mathcal{F}_p, \mathcal{F}_q),\ p \neq q$

How to find all “minimal” combinations of ap-hyperplanes that separate the sets $\mathcal{F}_1, \ldots, \mathcal{F}_s$?

Separation power of ap-hyperplanes

(Druhle et al., 2005)

- An ap-hyperplane has a supporting vector parallel to one axis
  - The label of the axis is the direction of the ap-hyperplane
- The separation power $S(\theta)$ of an ap-hyperplane θ describes the separated data sets
- Two ap-hyperplanes with a same direction and a same separation power are equivalent (thus defining equivalence classes of ap-hyperplanes)



$$dir(\theta_1) = dir(\theta_2) = 1$$

$$S(\theta_1) = S(\theta_2) = \{(1,2),(1,3)\}$$

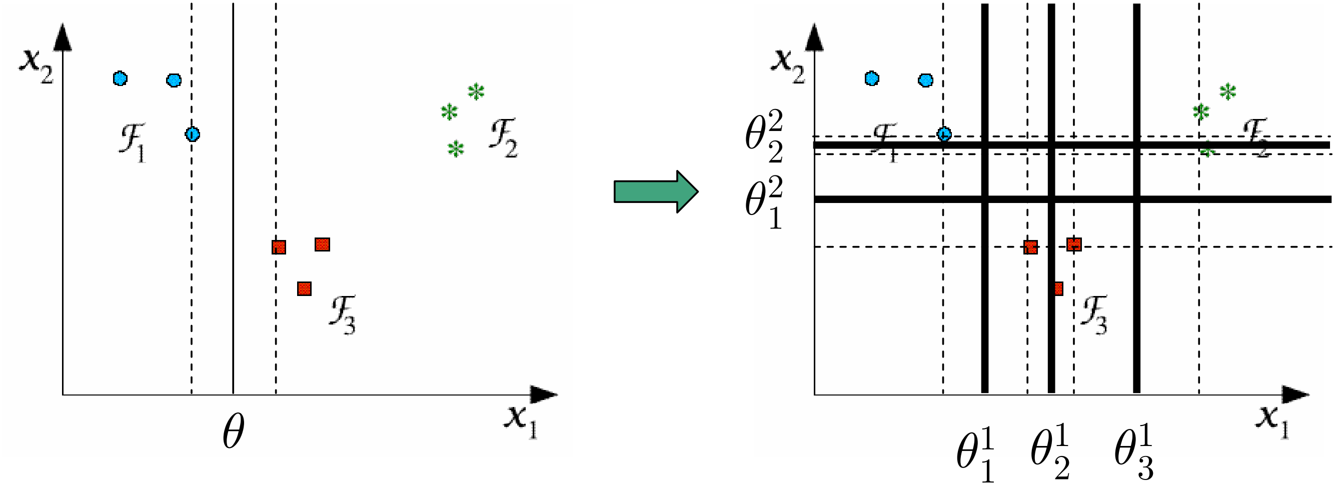
$$\Downarrow$$

$$\theta_1 \sim \theta_2$$

Cuts

For each class of equivalence, the ap-hyperplane that minimizes the empirical risk (i.e. that lies in the middle of the equivalence class) is a cut

The collection $\mathcal{C}^*$ of all cuts can be easily computed



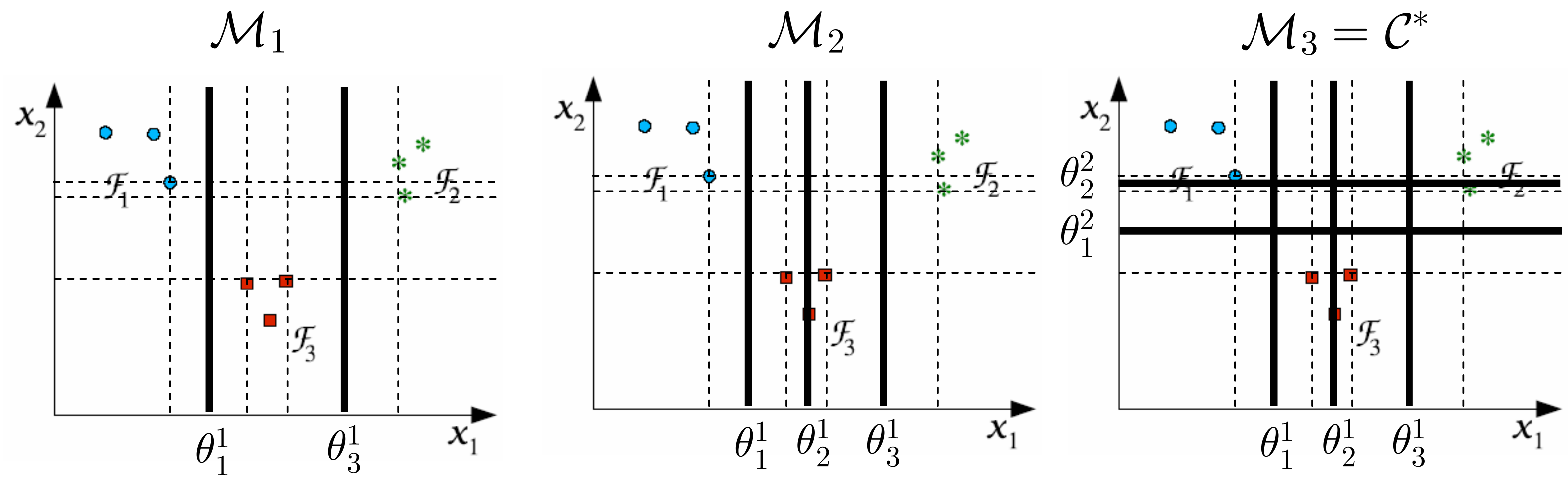
Standing assumption: all pairs of sets are separated by $\mathcal{C}^*$

$\mathcal{C}^*$ contains unnecessary cuts (i.e. unnecessary regulation circuits)

Occam's razor: find the *simplest* collections of cuts that separate the sets

Multicuts

A collection of cuts such that all pairs of sets are separated is a multicut



Rough idea: find all minimal multicuts by enumerating all multicuts

- combinatorial explosion !

Better ideas:

- remove cuts that are “redundant”
- find criteria for avoiding the enumeration of all multicuts

Multicut algorithm

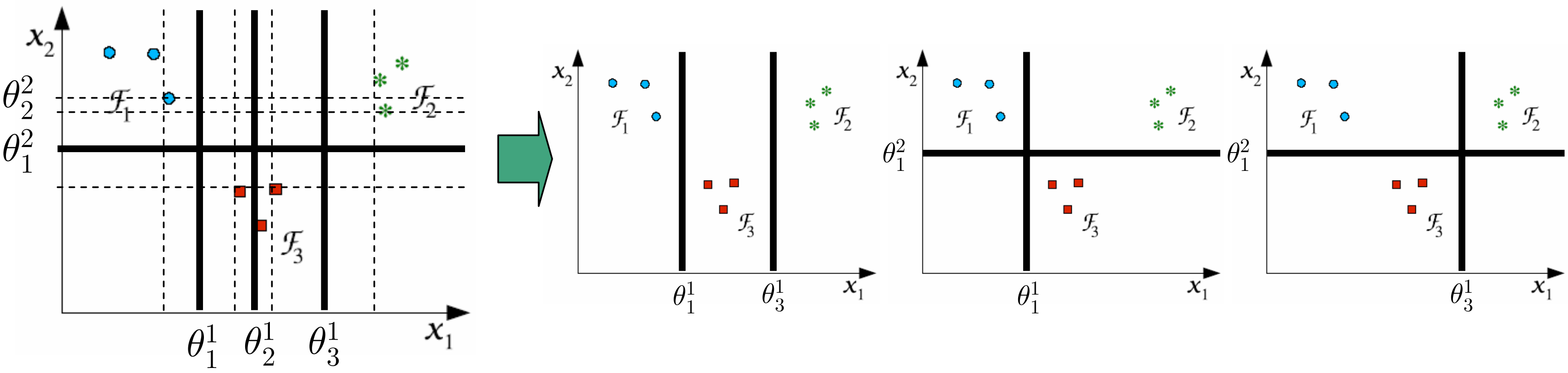
(Druhle et al., 2005)

- remove cuts that are “redundant”
- find criteria for avoiding the enumeration of all multicuts

How to do it ?

Mathematics: define partial order relations on cuts and multicuts and exploit the theory of POSETS.

Algorithms: branch-and-bound methods for computing all minimal multicuts

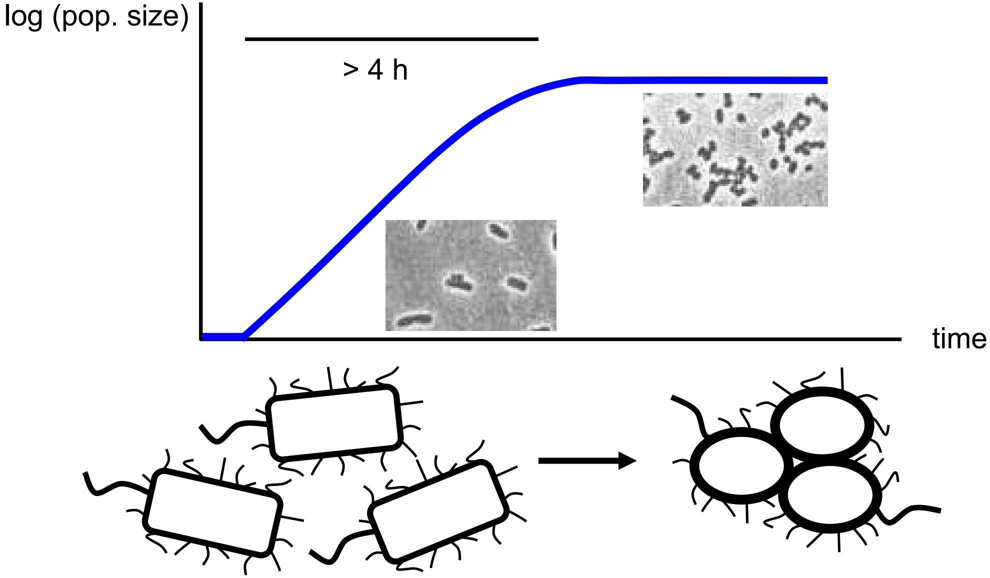


A case study

Identification of the GRN governing carbon starvation response of *E. coli*

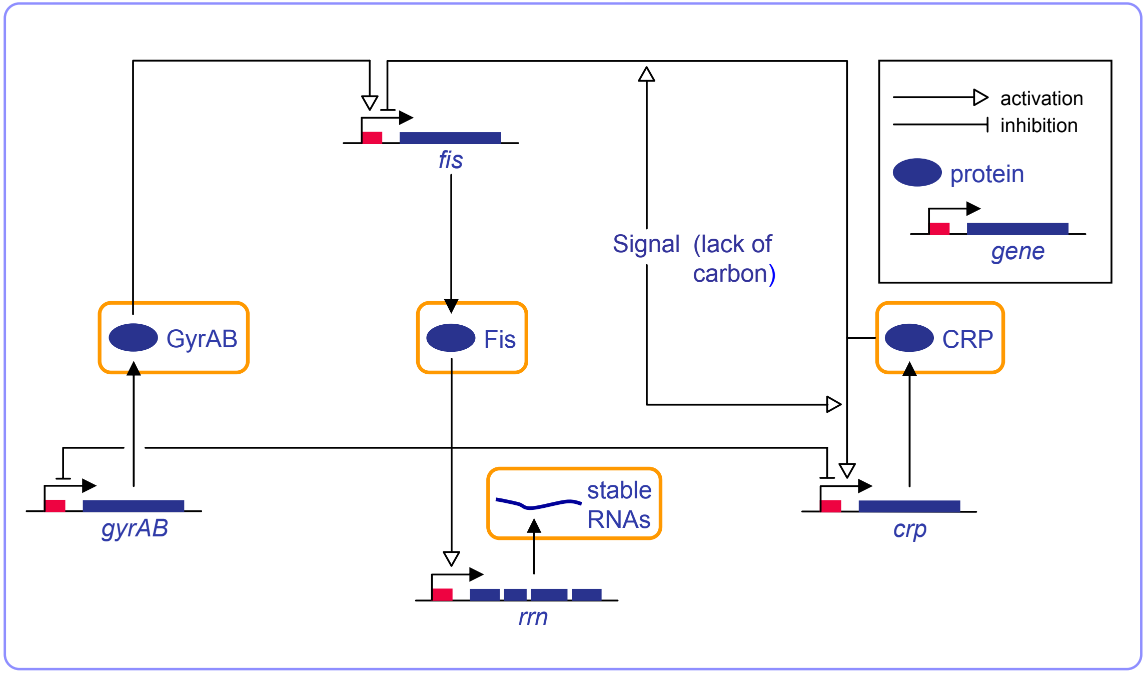
Transitions from exponential to stationary phase involve observable changes in:

- morphology,
- metabolism,
- gene expression,
- …



Simplified GRN

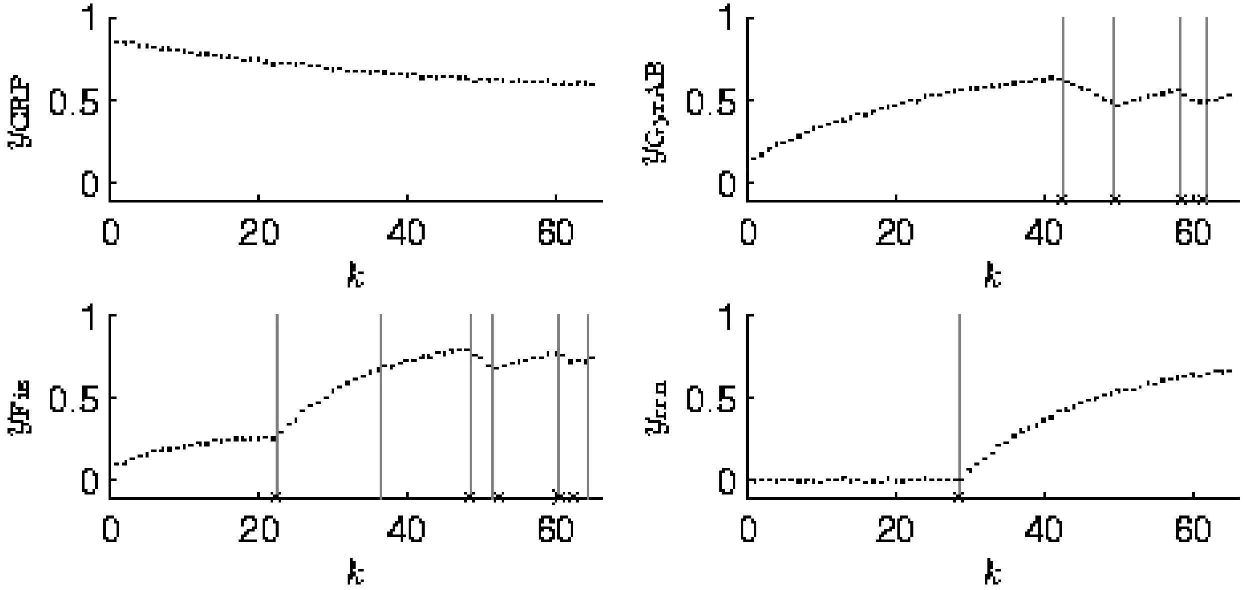
$$
\begin{aligned}
\dot{x}_{CRP} &= \kappa^0_{CRP} + \kappa^1_{CRP}\, s^-(x_{Fis}, \theta^1_{Fis})\, s^+(x_{CRP}, \theta^1_{CRP})\, s^+(x_S, \theta_S) - \gamma_{CRP}\, x_{CRP} \\
\dot{x}_{Fis} &= \kappa^1_{Fis}\, (1 - s^+(x_{CRP}, \theta^1_{CRP})\, s^+(x_S, \theta_S)) \\
&\quad + \kappa^2_{Fis}\, s^+(x_{GyrAB}, \theta_{GyrAB})\, (1 - s^+(x_{CRP}, \theta^1_{CRP})\, s^+(x_S, \theta_S)) - \gamma_{Fis}\, x_{Fis} \\
\dot{x}_{GyrAB} &= \kappa_{GyrAB}\, s^-(x_{Fis}, \theta^3_{Fis}) - \gamma_{GyrAB}\, x_{GyrAB} \\
\dot{x}_{rrn} &= \kappa_{rrn}\, s^+(x_{Fis}, \theta^2_{Fis}) - \gamma_{rrn}\, x_{rrn} \\
\dot{x}_S &= 0
\end{aligned}
$$



Switch detection

Data produced by an OE-PWA model (× = true switches)

- simulation of the transition stat. → exp. due to carbon upshift



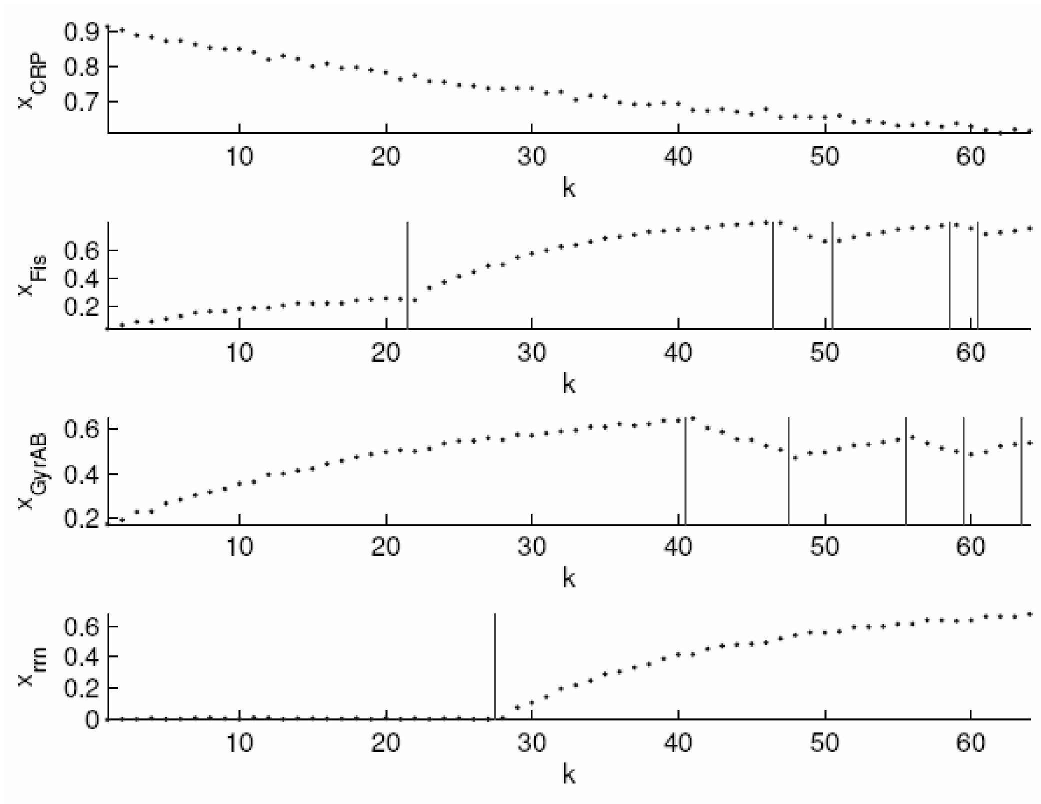
Vertical lines: switch detected by the algorithm based on nonlinear estimation

- all switches have been reconstructed
- one spurious switch in the profile of protein Fis

Reconstruction of switching thresholds

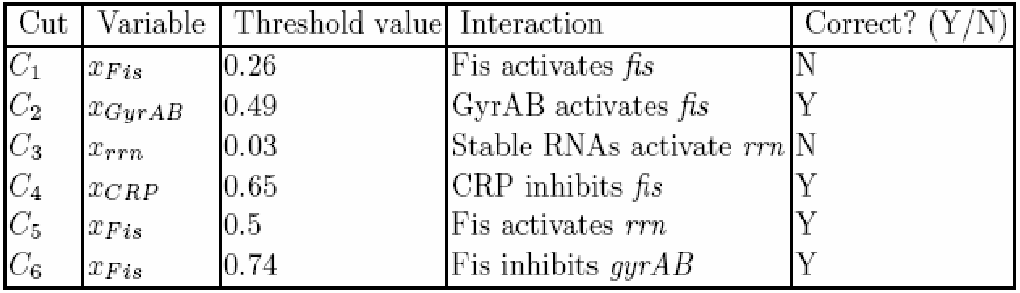
Data produced by a PWARX model (vertical lines = true switches)

- correct classification used for building the mode data sets $\mathcal{F}_1, \ldots, \mathcal{F}_s$



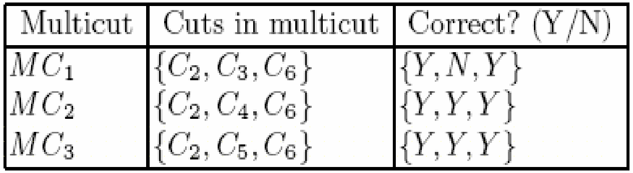
Reconstruction of switching thresholds

Non "redundant" cuts found by the algorithm:

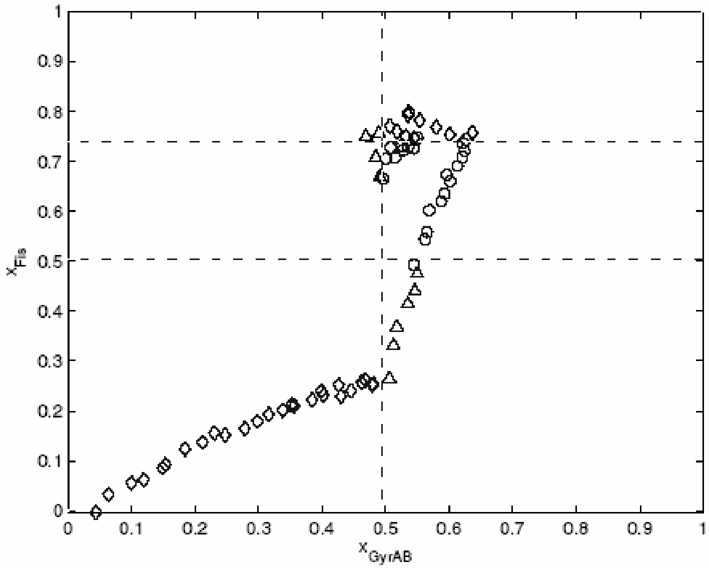


Cut	Variable	Threshold value	Interaction	Correct? (Y/N)
C_1	x_{Fis}	0.26	Fis activates *fis*	N
C_2	x_{GyrAB}	0.49	GyrAB activates *fis*	Y
C_3	x_{rrn}	0.03	Stable RNAs activate *rrn*	N
C_4	x_{CRP}	0.65	CRP inhibits *fis*	Y
C_5	x_{Fis}	0.5	Fis activates *rrn*	Y
C_6	x_{Fis}	0.74	Fis inhibits *gyrAB*	Y

Minimal multicuts found:

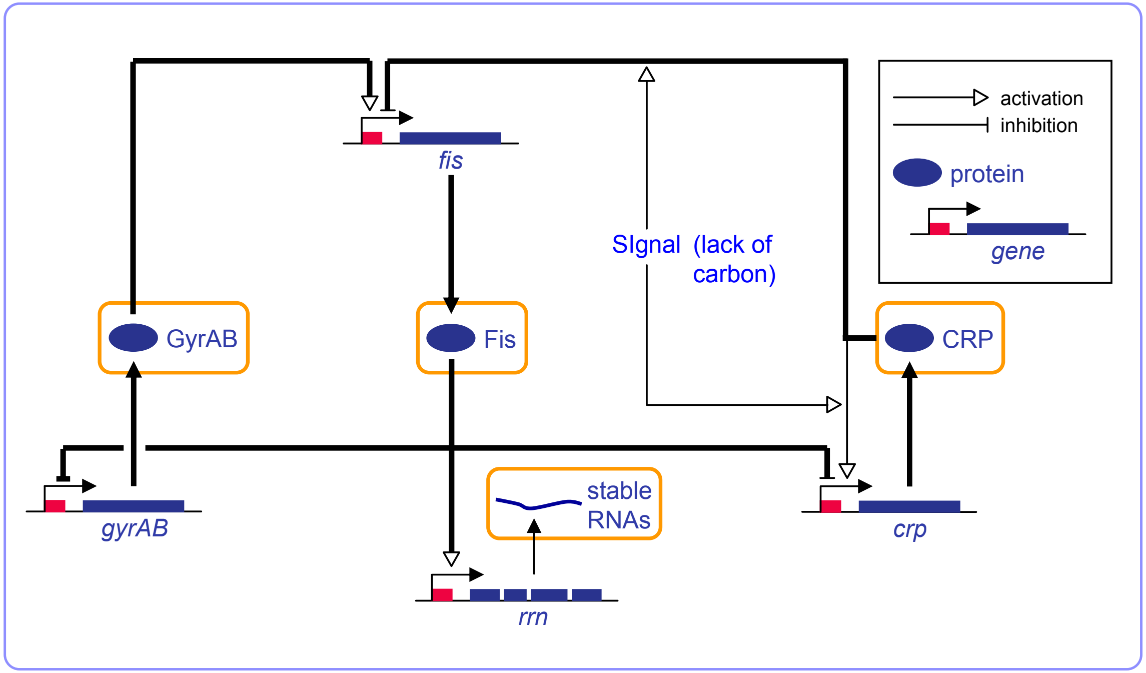


Multicut	Cuts in multicut	Correct? (Y/N)
MC_1	$\{C_2, C_3, C_6\}$	$\{Y, N, Y\}$
MC_2	$\{C_2, C_4, C_6\}$	$\{Y, Y, Y\}$
MC_3	$\{C_2, C_5, C_6\}$	$\{Y, Y, Y\}$



Reconstruction of switching thresholds

Merging the best minimal multicuts obtained on stat. → exp. and exp. → stat. data sets, only one interaction (autoactivation of CRP) has not been inferred



Conclusions

- Data-driven modeling of GRNs is a very active area of systems biology
  - Experimental techniques for obtaining accurate gene expression data are available

- Hybrid systems are appealing for modeling GRNs
  - compromise between linear and nonlinear models
  - they preserve the on/off behavior of genes

- Identification of PWA models of GRNs: exploit structure in order to
  - improve identification results
  - obtain multiple, biologically meaningful models

Current limitations of the proposed methods for switch detection and threshold reconstruction:

- absence of sliding-mode behaviors
- separability of mode data sets
- no capability of detecting “missing” genes