HAL
open science

Principal Process Analysis and reduction of biological models with order of magnitude

Stefano Casagranda, Jean-Luc Gouzé

To cite this version:

Stefano Casagranda, Jean-Luc Gouzé. Principal Process Analysis and reduction of biological models with order of magnitude. IFAC World Congress, International Federation of Automatic Control (IFAC). AUT., Jul 2017, Toulouse, France. 15968 p., 10.1016/j.ifacol.2017.08.2241 . hal-01529448

HAL Id: hal-01529448

https://inria.hal.science/hal-01529448

Submitted on 30 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Principal Process Analysis and reduction of biological models with order of magnitude

Stefano Casagranda * **Jean-Luc Gouzé** *

* *Université Côte d'Azur, Inria, INRA, CNRS, UPMC Univ Paris 06, France (e-mail: stefano.casagranda@inria.fr, jean-luc.gouze@inria.fr)*

Abstract: We present a simple method that allows to analyze the biological processes of a dynamical model and classify them. Along the system trajectories, we decompose the model into biological meaningful processes and then study their *activity* or *inactivity* during the time evolution of the system. The structure of the model is then reduced to the core mechanisms involving only the *active processes*. The initial conditions are supposed to lie in some rectangle, that could represent one order of magnitude for the variables. Keeping only the *active processes*, we obtain the principal processes in the rectangle and then in the adjacent rectangles where the trajectories may have a transition. Finally we obtain a partition of the space with a reduced model within each rectangle. We apply these techniques to a classical model of gene expression with protein and messenger RNA.

Keywords: Biological models, model reduction, dynamical systems, process analysis, gene expression.

1. INTRODUCTION

Mathematical models are relevant in understanding complex problems in numerous biology domains. A significant progress in modeling and parameter estimation approaches has lead to more and more complex kinetic models of cellular networks (see Bettenbrock et al. (2006) and Kuepfer et al. (2007) as examples). The large size and the complexity of these models cause problems for relating the global behavior of the system to the functioning of specific cellular processes (e.g RNA transcription, protein degradation, etc...). This information is crucial to understand what are the cellular processes that are giving a strong contribution for cellular growth and environmental adaptation and when they are at play or not. Efforts in this direction rely on the reduction of the system complexity, as reduced models can be more easily analyzed. Sensitivity analysis is often used to detect model parameters that are not influential for the system dynamics and that can be removed (i.e Degenring et al. (2004)). When different time scales are present, quasi-steady state approximation is often used to reduce the system dimension because a certain number of differential equations are substituted by algebraic equations. Nevertheless, the resulting differential algebraic system can be still difficult to analyze (Segel and Slemrod (1989)). Phase plane analysis instead allows analyzing the asymptotic behavior of kinetic models, but cannot apply to large systems (Khalil (1996),p. 21).
Other approaches simplify the function that describes the molecular processes: piece-wise affine differential equations approximate the sigmoidal functions describing the cooperativity in the regulation of gene expression, by step function. However, these simplifications are difficult to apply to model every types of network, and are restricted to models of gene expression (Baldazzi et al. (2010)).
We tackle the problem in this study with a mathematical and numerical approach, by decomposing the system into processes having a physical meaning, and considering the contribution of every process to the dynamics of the system: keeping only the relevant processes, based on a principal concept of *activity/inactivity*, we do not change the main structure of the system, but reduce it to its most important processes for a given interval of time. Because this approach is focused on analyzing the processes in a biological model and keeping only the principal ones to obtain a reduced model, we call it Principal Processes Analysis (PPA).
Our method is general and can be easily applied to any ODE model of biological system: for example, it has been recently used to reveal the correlation between C cycling and pesticide degradation in the detritusphere and fungal dynamics (Pagel et al. (2016)), or to reduce a dynamic metabolic model of lipid accumulation (Robles-Rodriguez et al. (2016)).
PPA shares some common features with the work of Apri et al. (2012), based on the exploration of the parameter space that leads to *admissible* system outputs. Their goal is to find the parameters that contribute the most to the system output and the ones that are not contained in this list are removed. The difference is that they search for important parameters and we look for important processes. Another work (Petzold and Zhu (1999)) aims at reducing chemical systems by removing the chemical species that contribute less to the model output. Their work is based on stoichiometric coefficients and chemical reactions and the problem is solved using optimization approaches. Although their goal is similar, their method is hence restricted to chemical kinetics (see also Bhattacharjee et al. (2003)). This work also has similarities with the work of one of the authors (Benoît and Gouzé (2009)), but this approach is not oriented toward reduction.
In our previous work (Casagranda et al. (2015)), we started

to develop our technique on two deterministic models. The first model described the *Drosophila* Circadian Rhythms dynamics (Leloup and Goldbeter (1998)): two reduced models have been created, one for the day period and the other for the night period, that were able to express the core of the system without changing significantly its dynamics. The second model (Kwang-Hyun et al. (2003)) described the inhibition of the activation of RAF by RKIP, which regulates the ERK signaling pathway: our reduction analysis showed that the most active process during the system dynamics influenced the most important variable found in Petrov et al. (2007).
The results were valid for a fixed set of initial values and parameter values, and the robustness of this approach needs to be tested. In this paper we focus on how the choice of the initial conditions affects the *activity* of the processes, resulting into a possible change of the reduced model.
Instead of a single initial point we consider the PPA on an entire set of possible initial values. For sake of simplicity and brevity, and because the orders of magnitude of the variables are very important in biological models, we consider initial conditions in rectangles representing one order of magnitude (e.g., the variables are between 1 and 10, or 10 and 100...) and we limit this first approach to the dimension two. It is however clear that it could be applied to any rectangular grid, and to any dimensions (but the notations would be more cumbersome). The plane (x_1,x_2) is therefore divided into a logarithmic grid, and we apply (under some assumptions concerning the monotonicity of the processes) our method by computing a maximal bound for the weight of each process within the rectangle. We only retain the *active* processes, having a dynamical weight higher that a fixed threshold δ.
The paper is organized as follows: in Section 2, we present the technique of reduction for a fixed initial condition, then compute the weights in the rectangle and finally within every rectangle of the space that can be reached from the initial rectangle. In Section 3 we present the gene model and in Sections 4 and 5 we apply the technique presented in the previous sections. The conclusions are presented in Section 6.

2. METHODOLOGY

2.1 Principal process analysis and model reduction

We briefly remind the process analysis and reduction method (see Casagranda et al. (2015)) for fixed initial conditions and parameter value. Consider the ODE system (1) that models a biological network (for example an intracellular network):

$$\dot{x} = f(x,p) \tag{1}$$

where $x = (x_1, x_2, .., x_n) \in \mathbb{R}^n$ is the vector of concentrations of the components, $x_0 = (x_{01}, x_{02}, ..., x_{0n}) \in \mathbb{R}^n$ is the vector of initial conditions and $p \in \mathbb{R}^b$ is the vector of parameters. It is possible to decompose each equation into a sum of biological processes:

$$\dot{x}_i = \sum_j f_{i,j}(x,p) \tag{2}$$

where $f_{i,j}(x,p)$ represents the jth process involved in the dynamical evolution of the ith variable of the system over a period of time [0,T].
In order to compare the influence of the different processes $f_{i,j}(x,p)$ in the evolution of each variable x_i, we associate a relative weight with each process to make it dimensionless:

$$W_{i,j}(t,p) = \frac{|f_{i,j}(x(t),p)|}{\sum_j |f_{i,j}(x(t),p)|} \tag{3}$$

where $0 \leq W_{i,j}(t,p) \leq 1$ and $\sum_j W_{i,j}(t,p) = 1$.

Definition: *Let the continuous function $f_{i,j}(x(t),p)$ be the jth process of $\dot{x}(t)_i$ for $t \in [0,T]$ and let the threshold $\delta \in$ [0,1]. We call a process $f_{i,j}(x(t),p)$ always inactive when $W_{i,j}(t,p) < \delta \ \forall \ t \in$ [0,T]. We call a process $f_{i,j}(x(t),p)$ inactive at time t when $W_{i,j}(t,p) < \delta$. We call a process $f_{ij}(x(t),p)$ active at time t when $W_{i,j}(t,p) \geq \delta$.*

The first step of the PPA is to identify the *always inactive* processes and delete them from the original system (1). The threshold value δ must be chosen between the range [0,1]: a low threshold avoids neglecting important processes.
The goal is to obtain a function $g(x^r)$ which approximates the function $f(x)$, that contains a minor number of processes. Let consider the ODE system $g(x^r)$ which approximates the system (1):

$$\dot{x}^r = g\left(x^r, p^r\right) \tag{4}$$

where $x^r = (x_1^r, x_2^r, .., x_n^r) \epsilon \mathbb{R}^n$ is the vector of concentration of the components, x_0 is the vector of their initial values and $p \epsilon \mathbb{R}^c$, where $c \leq b$ is the vector of the parameters.
The basic idea of the proposed model reduction method is based on the following classical theorem: if the vector fields of two systems are close ($f(x) \approx g(x)$), then the solutions of the original and approximated systems are close during some time interval under the assumptions on the Lispchitz conditions listed in Khalil (1996),p. 79,Th 2.5.
After having assigned dynamical weights to every process and a value to the threshold δ , we follow this rule to obtain $g(x^r)$:
if $W_{i,j}(x(t),p) < \delta \ \forall \ t \in$ [0,T] then $g_{i,j}(x(t),p) = 0$;
if not, $g_{i,j}(x(t),p) = f_{i,j}(x(t),p)$.
To test the quality of the reduced model $g(x^r)$, we numerically compute the global relative error between the original and reduced model for each variable. It is defined for the ith variable as:

$$e_i = \frac{\int |x_i(t) - x_i^r(t)| dt}{\int |x_i(t)| dt} \tag{5}$$

where $x_i(t)$ and $x_i^r(t)$ are respectively the solutions of the original and reduced systems. This method strongly depends on the initial condition.

2.2 Principal process analysis and model reduction based on initial conditions in a rectangle

To increase the robustness of the method, the initial condition is chosen in some region, then we compute if the *activity/inactivity* of the process $f_{i,j}$ - and consequently the reduced system $g(x^r)$ - changes.
We divide the variable space into rectangles, and then

apply the technique in each domain: for simplicity, we consider in this paper a system with two variables (x_1, x_2) and a logarithmic subdivision, corresponding to order of magnitude from the modeling point of view. Every point $\theta^{m,n} = (\theta_1^m, \theta_2^n)$ corresponds the value $(10^m, 10^n)$: for example the point $\theta^{2,0} = (10^2, 10^0) = (100, 1)$. We call $B_{m,n}$ the rectangle delimited by the four vertices $\theta^{m,n}$, $\theta^{m+1,n}$, $\theta^{m+1,n+1}$ and $\theta^{m,n+1}$ (shown in Figure 1): inside of it, every process $f_{i,j}(x,p)$ is limited horizontally $f_{i,j}(\theta_1^{m,n}, p) < f_{i,j}(x,p) < f_{i,j}(\theta_1^{m+1,n}, p)$ and vertically $f_{i,j}(\theta_2^{m,n}, p) < f_{i,j}(x,p) < f_{ij}(\theta_2^{m,n+1}, p)$.

To compute a global bound for the weights in the rectangle, we need the following assumption for the processes. Below, all the functions are supposed to be locally Lipschitz; by "fixed sign", we mean that the functions are either non-negative, or non-positive, or zero.

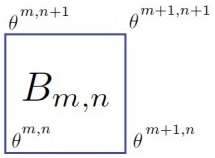

Fig. 1. A generic rectangle $B_{m,n}$ delimited by the vertices $\theta^{m,n}$, $\theta^{m+1,n}$, $\theta^{m+1,n}$, $\theta^{m+1,n+1}$.

Assumption: *$\partial f_{ij}/\partial x_k$ has a fixed sign in B_{mn}, $\forall i, k \in \{1, \ldots, n\}, \forall j$; moreover for given i and k, all the $\partial f_{ij}/\partial x_k$ have the same sign.*

In words, it means that all the processes for the velocity $\dot{x}_i$ have a derivative of a fixed sign with respect to any variable. This assumption is verified for many models.
Because of equation (2), it easily implies the following corollary.

Corollary: *The Jacobian matrix $J = Df(x,p)$ of the system (1) has a fixed sign inside the rectangle $B_{m,n}$.*

Remark that the Jacobian matrix is signed, but all the signs may be different (therefore the system is not monotone in the sense of conservation of partial order between trajectories). This assumption allows to study the behavior of the process $f_{i,j}(x,p)$ inside the full rectangle $B_{m,n}$, knowing only the behavior of the process at the vertices $\theta^{m,n}$, $\theta^{m+1,n}$, $\theta^{m+1,n+1}$ and $\theta^{m,n+1}$. Indeed, the monotonicity of each process with respect to any variable implies that:

Corollary: *In the rectangle $B_{m,n}$, each process $f_{i,j}$ takes its maximum and minimum on the vertices of the rectangle.*

We note $S_{i,j}^{m,n}$ the vertex of $B_{m,n}$ where the process f_{ij} is maximum, and $s_{i,j}^{m,n}$ the vertex of $B_{m,n}$ where the process f_{ij} is minimum in $B_{m,n}$.

Inside $B_{m,n}$, a worst-case version of the general weight (3) is

$$WW_{i,j}^{B_{m,n}}(p) = \frac{|f_{i,j}(S_{i,j}^{m,n}, p)|}{\sum_j |f_{i,j}(s_{i,j}^{m,n}, p)|} \tag{6}$$

and normalizing these weights to proportions summing to one we obtain:

$$W_{i,j}^{B_{m,n}}(p) = \frac{|f_{i,j}(S_{i,j}^{m,n}, p)|}{\sum_j |f_{i,j}(S_{i,j}^{m,n}, p)|} \tag{7}$$

The reduction method in $B_{m,n}$ is now similar to the previous one: if this weight is smaller than some threshold δ, then the process is considered as *inactive* in $B_{m,n}$, and discarded. A reduced model is obtained within each domain $B_{m,n}$ by keeping the principal processes.

2.3 Possible transitions between domains

Furthermore, our assumption has strong consequences concerning the possible transitions between rectangles. These transitions are conditioned by the vector field on the boundary (the edges) of each rectangle $B_{m,n}$. This analysis is valid for the full or the reduced model.

Proposition: *For $i \in \{1, 2, ..., n\}$, if $\dot{x}_i$ is positive (resp. negative) on two adjacent vertices, then it is positive (resp. negative) on the edge between these two vertices. If the signs are opposite, then the vector field will cancel somewhere on the edge.*

Proof: By Corollary (2), the Jacobian matrix is signed, and therefore each component of the vector field is increasing or decreasing along an edge.

Therefore, knowing the values of the processes at the vertices of the rectangle, allows to study the possible transitions of solutions $x(t)$ from a rectangle $B_{m,n}$ into another adjacent rectangle or into the same rectangle.
As an example, in Figure 2 two different situations are illustrated. In Fig. 3.A, the solutions of the system x moves in the rectangles $B_{m+1,n}$ and $B_{m,n+1}$. In Fig. 3.B, the solutions stay in the rectangle $B_{m,n}$, which is invariant.

Therefore, having the reduced model in some rectangle $B_{m,n}$, the vector field on the vertices gives the possible transitions toward adjacent rectangles. Then the model is reduced in these rectangles. Finally a graph of transition is obtained between rectangles, each rectangle having a reduced model. The biologist may follow the possible sequence of reduced models with respect to the different orders of magnitude. This sequence is not deterministic because most of the time a rectangle has transitions in several rectangles.
Moreover, from the sequence of reduced models, if some process is *always inactive* for every rectangle, it is possible to obtain a global reduced model valid on the whole pathway of rectangles.

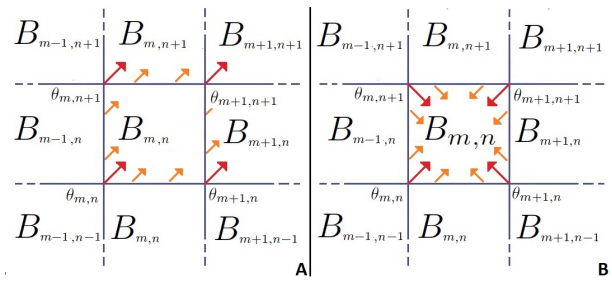

Fig. 2. In Figure A (left) the vector field leads the solutions of the system to move from $B_{m,n}$ to the adjacent rectangles, In Figure B (right) the vector field leads the solutions to stay inside the rectangle $B_{m,n}$.

3. THE GENE EXPRESSION MODEL

We apply this technique on a classical deterministic model in which the protein P inhibits its own mRNA (Bernot et al. (2013), p.57).
The amount of mRNA produced (variable M) depends on its basal activity κ_1 and on the transcription activity, based in turn on the concentration of its DNA sites bound to the repressor P and on the amount of the free DNA sites, $\kappa_2 \frac{\alpha_P^m}{\alpha_P^m + P^m}$. The mRNA can degrade with a degradation term γ_M and be diluted due to growth rate (μ). The translation process leads the production of the protein P and it is designed as a linear function of the mRNA ($\kappa_3 M$) . The protein can also degrade with a degradation term γ_P and be diluted due to growth rate (μ).

$$\frac{d}{dt} M = \kappa_1 + \kappa_2 \frac{\alpha_P^m}{\alpha_P^m + P^m} - (\gamma_M + \mu) M \tag{8}$$

$$\frac{d}{dt} P = \kappa_3 M - (\gamma_P + \mu) P . \tag{9}$$

In Table 1 are presented the model parameters with their units: we used the information contained in database for biological numbers, BioNumbers, to give to model parameters reasonable values (Milo et al. (2010)).

Table 1. Parameters

Parameter	Value	Unit
κ_1	0.000925	$\mu M min^{-1}$
κ_2	0.00185	$\mu M min^{-1}$
κ_3	1.39	–
γ_M	0.1733	min^{-1}
γ_P	0.0023	min^{-1}
m	2	–
α_P	20	μM
μ	0.0166	$\mu M min^{-1}$

4. MODEL REDUCTION FROM AN INITIAL CONDITION

As a first step, we decompose the ODE system (8-10) in the following processes. The mRNA derivative ($\frac{d}{dt}M$) can be divided into its basal activity process $f_{1,1} = \kappa_1$, into transcription process $f_{1,2} = \kappa_2 \frac{\alpha_P^m}{\alpha_P^m + P^m}$, into degradation process $f_{1,3} = \gamma_M M$ and into dilution process $f_{1,4} = \mu M$; The protein derivative ($\frac{d}{dt}P$) can be divided into translation process $f_{2,1} = \kappa_3 M$, into degradation process $f_{2,2} = \gamma_P P$ and into dilution process $f_{2,3} = \mu P$.
We calculate the process weights of the system using formula (3), having a initial conditions vector: $x_0 = [\theta_1^0, \theta_2^0]$. In Figure 3 are shown the plots of the weights of the processes for a fix threshold of $\delta = 0.2$.
From this analysis, the processes resulting *always inactive* are $f_{1,1}$, $f_{1,3}$, $f_{2,2}$. The new system $g(x^r)$ is:

$$\frac{d}{dt} M^r = \kappa_2 \frac{\alpha_P^m}{\alpha_P^m + (P^r)^m} - \gamma_M M^r \tag{10}$$

$$\frac{d}{dt} P^r = \kappa_3 M^r - \mu P^r . \tag{11}$$

The global relative errors between the solution of the variable P in the original system $f(x)$ and the reduce one $g(x^r)$ are shown in Table 2.

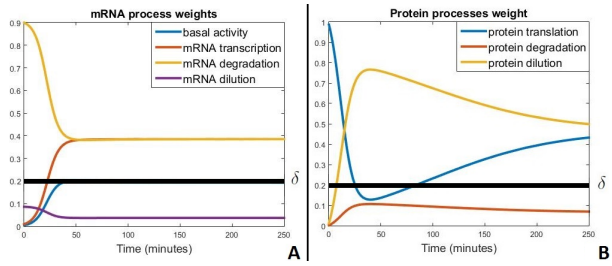

Fig. 3. A. The evolution in time of mRNA process weights: the basal activity process and the degradation process are *always inactive* because the dynamics are always under the threshold δ. B. The evolution in time of protein process weights: the dilution process is *always inactive* because the dynamic is always under the threshold δ.

Table 2. Global Relative Errors

Variable	values
e_M	0.14
e_P	0.10

5. MODEL REDUCTION IN A RECTANGLE

We extend our method to the entire rectangle $B_{0,0}$ that has the vertices ($\theta^{0,0}$, $\theta^{0,1}$, $\theta^{1,1}$, $\theta^{1,0}$). We first verify the assumption on the monotonicity of the processes and the Jacobian matrix, written as:

$$J = \begin{bmatrix} \frac{df_1}{dM} & \frac{df_1}{dP} \\ \frac{df_2}{dM} & \frac{df_2}{dP} \end{bmatrix} = \begin{bmatrix} -(\gamma_M + \mu) & -\kappa_2 \frac{\alpha_P^m P^m m}{(\alpha_P^m + P^m)^2 P} \\ \kappa_3 & -(\gamma_P + \mu) \end{bmatrix}$$

We compute the vector field for the rectangle $B_{0,0}$, $\theta_1^0 < x_1 < \theta_1^1$ and $\theta_2^0 < x_2 < \theta_2^1$ at the 4 vertices and because of the monotonicity of the Jacobian matrix, we can deduce the behavior of the processes on the edges of the rectangle. The result is presented in Figure 4. Based on the direction of the arrows, the solutions move to the rectangles $B_{-1,0}$ and $B_{0,1}$.
Using the formula (7), it is possible to compute the weight for the entire rectangle $B_{0,0}$, based on the worst case. In Table 3, for the rectangle $B_{0,0}$ we show the maximum value that the process $f_{i,j}(x,p)$ can reach and its weights: setting the value of the threshold δ at 0.2, we can neglect the processes $f_{1,1}$, $f_{1,2}$, $f_{1,4}$, $f_{2,2}$, $f_{2,3}$.

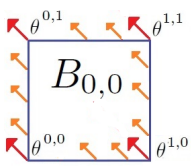

Fig. 4. Vector field on the edges. The solutions are moving into the adjacent rectangles $B_{-1,0}$ and $B_{0,1}$.

The valid sub-model $g(x^r)$ for $B_{0,0}$ is:

$$\frac{d}{dt} M^r = -\gamma_M M^r \tag{12}$$

$$\frac{d}{dt} P^r = \kappa_3 M^r \tag{13}$$

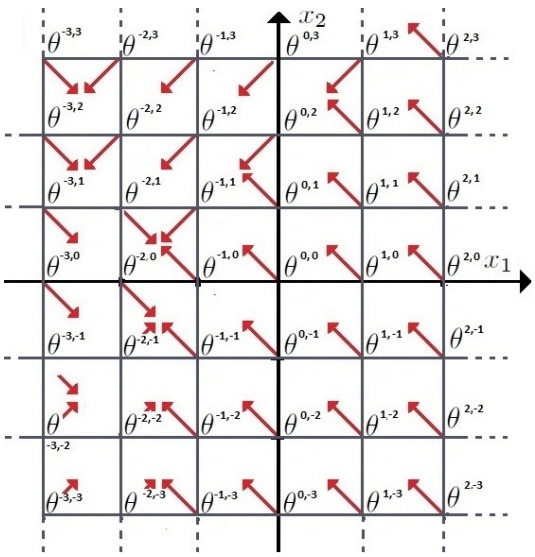

Fig. 5. Vector field on the edges in the plane.

Table 3. Processes in $B_{0,0}$

Process	Max. Value (μM)	Weight
κ_1	0.000925	0.00048638
$\kappa_2 \frac{\alpha^m}{\alpha^m + P^m}$	0.0018455	0.00097
$\gamma_M M$	1.733	0.91125
μM	0.166	0.087287
$\kappa_3 M$	13.9	0.9866
$\gamma_P P$	0.023	0.0016
μP	0.166	0.0118

The sub-model (12-13) is only valid in the rectangle $B_{0,0}$. To study the dynamics of the process weights over the whole time $[0, T]$ as we did for $\theta_{0,0}$ in Figure 3, we need to know the pathway of the solutions x in the different rectangles.

Extending the PPA as we did in Figure 4 to the full domain, we obtain the result shown in Figure 5: from any initial value x_0 the solutions are moving into the final rectangles $B_{-2,0}$ $B_{-2,-1}$. Starting from an initial value inside the rectangle $B_{0,0}$ we can have different solution pathways: the solutions can move into rectangle $B_{-1,0}$ or $B_{0,1}$, then from $B_{-1,0}$ they can move into $B_{-2,0}$ or $B_{-1,1}$ and from $B_{0,1}$ to $B_{0,2}$ or $B_{-1,1}$. Every pathway eventually ends in the space occupied by the rectangles $B_{-2,0}$ and $B_{-2,-1}$.

To explain the application of our technique, we perform it on one of the possible pathways that starts from the rectangle $B_{0,0}$: $B_{0,0} \Longrightarrow B_{0,1} \Longrightarrow B_{-1,1} \Longrightarrow B_{-2,1} \Longrightarrow B_{-3,1} \Longrightarrow B_{-3,0} \Longrightarrow B_{-2,0}$. The process weights of mRNA and protein, using formula (6) in each rectangle (or region) are plotted in Figure 6. Neglecting the *always inactive processes* we obtain the global reduced model (14-15).

$$\frac{d}{dt} M^r = \kappa_2 \frac{\alpha_P^m}{\alpha_P^m + (P^r)^m} - \gamma_M \, M^r \tag{14}$$

$$\frac{d}{dt} P^r = \kappa_3 \, M^r - \mu \, P^r \, . \tag{15}$$

The reduced model has the same structure of (10-11) with the difference that the first reduced model describes the dynamics of the system starting from a single initial value $\theta_{0,0}$ while the second one describes the dynamics of the system starting from any point of an entire region of initial values $B_{0,0}$ - in which the point $\theta_{0,0}$ is included - and follows a pathway till it arrives in the rectangle $B_{-2,0}$ which contains the steady-state of the solutions x of the original system (8).

Figure 7 represents a graphic way to obtain quickly the knowledge of the *activity/inactivity* of each process (black means *active* and white *inactive*) in each rectangle. Regarding mRNA processes it is possible to see that the basal activity is active when the mRNA has very low values and the protein has low values; the transcription is active only for small concentration of M and high concentration of P while the degradation is always an active process, in every rectangle. Regarding the protein processes it is possible to see that while the degradation process is always inactive in every rectangle, the translation process is active in the rectangle where the protein has a small concentration and the dilution is active when the protein has an high concentration. This information is very useful for the biological analysis of the system.

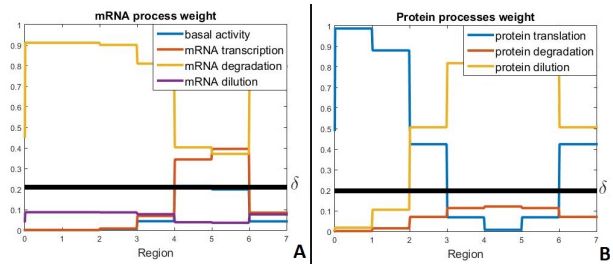

Fig. 6. A. The evolution in each rectangle of the mRNA process weights: the basal activity and mRNA dilution are *always inactive* because the dynamic is always under the threshold δ. The regions correspond respectively to the rectangles $B_{0,0}, B_{0,1}, B_{-1,1}, B_{-2,1}, B_{-3,1}, B_{-3,0}, B_{-2,0}$. B. The evolution in each rectangle of the protein process weights: the degradation is *always inactive* because the dynamic is always under the threshold δ. The regions correspond respectively to the rectangles $B_{0,0}, B_{0,1}, B_{-1,1}, B_{-2,1}, B_{-3,1}, B_{-3,0}, B_{-2,0}$

6. CONCLUSION

In this paper we proved the robustness of our technique in relation with a variation of the initial conditions of the system (8). In fact we have obtained a reduced model (10-11) applying our method to the original model that had an initial value $x_0 = [1, 1]$ and then we have obtained the same reduced model (14-15) choosing a space $B_{0,0}$ of initial values that contains $x_0 = [1, 1]$, a range of one order of magnitude in each coordinate and that follows a pathway close to the evolution of the system (8-9), starting from x_0 and ending in the steady state point x^*.

Furthermore, in every rectangle $B_{m,n}$ - that represents a different order of magnitude of the system - we obtain a meaningful reduced model in which we can obtain the knowledge of the *activity/inactivity* of each process as we presented in Figure 7. The biological interpretation of this table can be very fruitful. We used a grid, in which every boundary differs of one order of magnitude in relation to the previous one: a different grid can be chosen.

At first we have tested the robustness of our method on

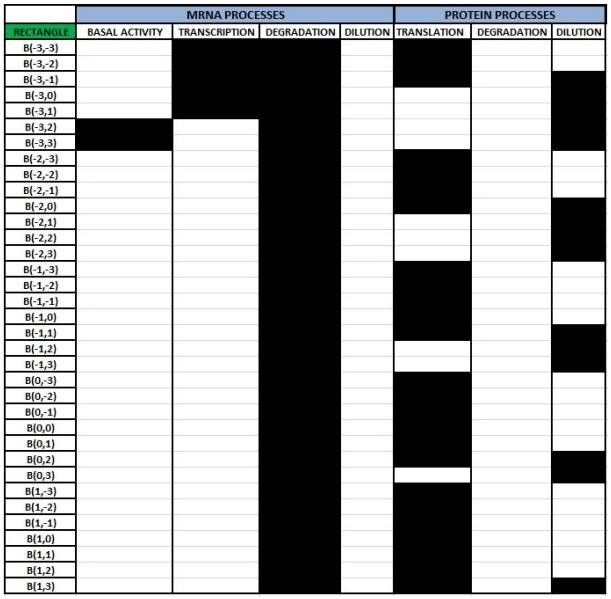

Fig. 7. The *activity/inactivity* of every process in each rectangle: black means that the process is *active* in that rectangle, white means that it is *inactive*

a model of two dimensions for simplicity reasons and to describe easily the applications: a future work will verify the robustness of our method on models of higher dimensions. Furthermore we can extend our analysis applying the same method as in Section 5 to the model parameters. Finally a further method of reduction could be applied: in Section 2.3 we consider that the possible reduction is done independently for each component of the vector field. We could also consider a more global reduction on the sum of the components.

ACKNOWLEDGEMENTS

We acknowledge *Conseil Régional PACA* and the *Investissements d'Avenir Bio-informatique programme* under project RESET (ANR-11-BINF-0005) for funding the PhD thesis of S.Casagranda. We thank the research program Labex SIGNALIFE (ANR-11-LABX-0028-01).
We also thank Delphine Ropers (Inria, IBIS) for her help with the choice of parameter values of the model presented in Section 3.

REFERENCES

Apri, M., de Gee, M., and Molenaar, J. (2012). Complexity reduction preserving dynamical behavior of biochemical networks. *Journal of theoretical biology*, 304, 16–26.

Baldazzi, V., Ropers, D., Markowicz, Y., Kahn, D., Geiselmann, J., and De Jong, H. (2010). The carbon assimilation network in *Escherichia coli* is densely connected and largely sign-determined by directions of metabolic fluxes. *PLoS Computational Biology*, 6(6), e1000812.

Benoît, E. and Gouzé, J.L. (2009). An algorithmic approach to orders of magnitude in a biochemical system. In *Positive Systems*, 233–241. Editors: Rafael Bru, Sergio Romero-Vivò. Springer.

Bernot, G., Comet, J.P., Richard, A., Chaves, M., Gouzé, J.L., and Dayan, F. (2013). Modeling and analysis of gene regulatory networks. In *Modeling in Computational Biology and Biomedicine*, 47–80. Editors: Frédéric Cazals and Pierre Kornprobst. Springer.

Bettenbrock, K., Fischer, S., Kremling, A., Jahreis, K., Sauter, T., and Gilles, E.D. (2006). A quantitative approach to catabolite repression in *Escherichia coli*. *Journal of Biological Chemistry*, 281(5), 2578–2584.

Bhattacharjee, B., Schwer, D.A., Barton, P.I., and Green, W.H. (2003). Optimally-reduced kinetic models: reaction elimination in large-scale kinetic mechanisms. *Combustion and Flame*, 135(3), 191–208.

Casagranda, S., Ropers, D., and Gouzé, J.L. (2015). Model reduction and process analysis of biological models. In *Control and Automation (MED), 2015 23rd Mediterranean Conference on*, 1132–1139. IEEE.

Degenring, D., Froemel, C., Dikta, G., and Takors, R. (2004). Sensitivity analysis for the reduction of complex metabolism models. *Journal of Process Control*, 14(7), 729–745.

Khalil, H.K. (1996). *Nonlinear Systems*. Prentice-Hall, New Jersey, second edition.

Kuepfer, L., Peter, M., Sauer, U., and Stelling, J. (2007). Ensemble modeling for analysis of cell signaling dynamics. *Nature biotechnology*, 25(9), 1001–1006.

Kwang-Hyun, C., Sung-Young, S., Hyun-Woo, K., Wolkenhauer, O., McFerran, B., and Kolch, W. (2003). Mathematical modeling of the influence of RKIP on the ERK signaling pathway. In *International Conference on Computational Methods in Systems Biology*, 127–141. Springer.

Leloup, J.C. and Goldbeter, A. (1998). A model for circadian rhythms in *Drosophila* incorporating the formation of a complex between the PER and TIM proteins. *Journal of biological rhythms*, 13(1), 70–87.

Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). Bionumbersthe database of key numbers in molecular and cell biology. *Nucleic acids research*, 38(suppl 1), D750–D753.

Pagel, H., Poll, C., Ingwersen, J., Kandeler, E., and Streck, T. (2016). Modeling coupled pesticide degradation and organic matter turnover: From gene abundance to process rates. *Soil Biology and Biochemistry*, 103, 349–364.

Petrov, V., Nikolova, E., and Wolkenhauer, O. (2007). Reduction of nonlinear dynamic systems with an application to signal transduction pathways. *IET systems biology*, 1(1), 2–9.

Petzold, L. and Zhu, W. (1999). Model reduction for chemical kinetics: An optimization approach. *AIChE Journal*, 45(4), 869–886.

Robles-Rodriguez, C., Bideaux, C., Guillouet, S., Gorret, N., Roux, G., Molina-Jouve, C., and Aceves-Lara, C. (2016). Multi-objective particle swarm optimization (mopso) of lipid accumulation in fed-batch cultures. In *Control and Automation (MED), 2016 24th Mediterranean Conference on*, 979–984. IEEE.

Segel, L.A. and Slemrod, M. (1989). The quasi-steady-state assumption: a case study in perturbation. *SIAM review*, 31(3), 446–477.