Model reduction and process analysis of biological models

Stefano Casagranda, Delphine Ropers, Jean-Luc Gouzé

Model reduction and process analysis of biological models

Stefano Casagranda[1], Delphine Ropers[2] and Jean-Luc Gouzé[3]

Abstract— **Understanding the dynamical behavior of biological networks is complicated due to their large number of components and interactions. We present a method to analyse key processes for the system behavior, based on the *a priori* knowledge of the system trajectory and the simplification of mathematical models of these networks. The method consists of the model decomposition into biologically meaningful processes, whose *activity* or *inactivity* is evaluated during the time evolution of the system. The structure of the model is reduced to the core mechanisms involving *active* processes only. We assess the quality of the reduction by means of global relative errors and apply our method to two models of the circadian rhythm in Drosophila and the influence of RKIP on the ERK signaling pathway.**

Keywords: Dynamical Systems, Biological Networks, Model Reduction, Process Analysis

I. INTRODUCTION

Mathematical modeling and simulation tools are determinant in understanding how adaptation of living organisms to environmental cues results from large networks of metabolites, RNAs, proteins, and their mutual interactions. Larger and larger kinetic models of cellular networks are nowadays published, as a result of decades of work in biology, recent advances in high throughput technologies, and progress in modeling and parameter estimation approaches (see [1] and [2] for an example).

The large size of these models and their non linearity due to complex feedback loops make their dynamical analysis rather difficult. More specifically, it is extremely difficult to relate the global behaviour of the system to the functioning of specific cellular processes (e.g. RNA transcription, protein phosphorylation, or complex formation), while this gain of knowledge is crucial to identify what are the key cellular processes for the environmental adaptation and when they are at play.

Efforts in this direction rely on the reduction of the system complexity, as reduced models can be more easily analysed. Beside testing model robustness, sensitivity analysis is hence used to detect model parameters that are not influential for the system dynamics and that can be removed (i.e. [3]). When different time scales are present in a model, quasi-steady-state approximation is often applied to reduce the

system dimension and replace the fast variables by algebraic expressions. Nevertheless, the resulting differential algebraic system can be still complex to analyse [4]. Finally, phase plane analysis allows analysing the asymptotic behaviour of kinetic models, but its use is restricted to small size systems [5, p. 21]. Alternative approaches consist in simplifying the mathematical functions describing the molecular processes. For instance, piece-wise affine differential equations approximate by step functions the sigmoidal functions describing cooperativity in the regulation of gene expression. The dynamics of this simplified system can be easily analysed by means of state transition graphs [6]. However, these simplifications are generally restricted to models of gene expression and are more difficult to apply to model other types of network [7].

We tackle the problem in this study, by developing a mathematical and numerical approach for analysing the contribution of cellular processes to the dynamics of an ordinary differential equation system. The method allows i) to identify, at different times of the dynamics, the mechanisms that are *active* and contribute most to the system dynamics and ii) to reduce the kinetic model into a series of new models reduced to the core mechanisms, neglecting the *inactive* processes. The method is general and can be easily applied to any ODE model of biological system.

To some extent, our work is similar to the model reduction approach of L. Petzold and W. Zhu [8] applied to chemical kinetics, which consists in decomposing ODE models into a sum of elementary reactions and finding simplified models with a reduced number of reactions. Removing part of the network is also considered by Apri *et al.* in [9], where parameters that are not important for the system dynamics are removed from the model.

We apply our method to two models with different dynamical behaviors. In the model of the circadian rhythm in Drosophila [10], the system dynamics reaches a regime of sustained periodic oscillations, whereas this dynamics is simpler in the model of the influence of Raf Kinase Inhibitor Protein on the Extra-cellular Signal Regulated Kinase [11]. The solutions quickly reach a steady-state. Section 2 of this paper describes the methodology of the model simplification. The method is applied to the circadian and ERK models in Section 3 and 4, respectively. We conclude in Section 5, by giving some future perspectives.

II. METHODOLOGY

We evaluate the different processes along the trajectory of the model, in order to study their relative influence during the evolution of the system. As such, this technique takes

[1] S. Casagranda: BIOCORE INRIA, Unité de recherche Sophia Antipolis, 2004 route des Lucioles, BP 93, 06902 Sophia Antipolis, France `stefano.casagranda@inria.fr`

[2] D. Ropers: IBIS INRIA, Unité de recherche Rhône-Alpes, 655 avenue de l'Europe, Montbonnot, 38334 Saint Ismier Cedex, France `delphine.ropers@inria.fr`

[3] J.-L. Gouzé: BIOCORE INRIA, Unité de recherche Sophia Antipolis, 2004 route des Lucioles, BP 93, 06902 Sophia Antipolis, France `jean-luc.gouze@inria.fr`

naturally into account the order of magnitude of the variables. Consider a general ODE model of biological network:

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

and an ODE system which approximates the previous model:

$$\dot{x}^r = g(x^r, p^r) \qquad (2)$$

where $x = (x_1, x_2, ..., x_n) \, \epsilon \, R^n$ and $x^r = (x_1^r, x_2^r, ..., x_n^r)$ $\epsilon \, R^n$ represent, respectively, the vectors of concentration of components in the original and reduced models. $p \, \epsilon \, R^m$ and $p^r \, \epsilon \, R^k$ with $k \leq m$ are respectively the vectors of parameters of the original and reduced models, where each parameter has a precise and known value. The vector of initial conditions is $x_0 = x_0^r$, where $x_0 \, \epsilon \, R^n$ is also known. The time interval on which we want to study the processes is $[0, T]$ (it can be either a short or a long period).

The basic idea of our method is based on the following 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 Lipschitz conditions listed in [5, p. 79,Th. 2.5]. We first identify the processes in the model equations.

$$\dot{x}_i = \sum_j f_{ij}(x, p) \qquad (3)$$

where f_{ij} represents the jth process involved in the evolution of the ith coordinate where $i = 1, ..., n$. Generally a biological model is already decomposed in terms of a linear combination of processes: they evolve dynamically over time and contribute to the time evolution of one or more variables. Let us take as an example the 8th coordinate of the Goldbeter and Leloup's model [10] (see Section 3 and the model equations in Appendix).

$$\frac{dT_2}{dt} = V_{3T}\frac{T_1}{K_{3T} + T_1} - V_{4T}\frac{T_2}{K_{4T} + T_2} - k_3 P_2 T_2 \qquad (4)$$

$$+ k_4 C - v_{dt}\frac{T_2}{K_{dT} + T_2} - k_d T_2$$

The equation represents the variation of the concentration of the double phosphorylated *timeless* protein. It contains six different processes, each of which with a specific biological meaning. They can be positive or negative.

$$\dot{x}_8 = f_{8,1} - f_{8,2} - f_{8,3} + f_{8,4} - f_{8,5} - f_{8,6} \qquad (5)$$

where $f_{8,1} = V_{3T}\frac{T_1}{K_{3T}+T_1}, ..., f_{8,6} = k_d T_2$.

In order to compare the influence of the different processes f_{ij} in the time evolution of each variable x_i, we have to find a criterion to compare the processes between themselves. It is possible, for example, to compare their absolute function ($|f_{ij}(x, p)|$), to scale it by the ith initial condition ($\frac{|f_{ij}(x(t),p)|}{x_{0i}}$) or to scale it by the solution of the ith ODE ($\frac{|f_{ij}(x(t),p)|}{x(t)_i}$). In this work we choose to associate a relative weight to each process with the purpose to make it dimensionless:

$$W_{ij}(t, p) = \frac{|f_{ij}(x(t), p)|}{\sum_j |f_{ij}(x(t), p)|} \qquad (6)$$

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

Definition: *Let the continuous function $f_{ij}(x(t), p)$ be the jth process of $\dot{x}(t)_i$ in $t \, \epsilon \, [0, T]$ and let the threshold $\theta \subset R \; \theta \, \epsilon \, [0,1]$.*
We call a process $f_{ij}(x(t), p)$ always inactive when $W_{ij}(t, p) < \theta \; \forall \; t \, \epsilon \, [0,T]$.
We call a process $f_{ij}(x(t), p)$ inactive at time t when $W_{ij}(t, p) < \theta$.
We call a process $f_{ij}(x(t), p)$ active at time t when $W_{ij}(t, p) \geq \theta$.
The switching time for a process $f_{ij}(x(t), p)$ is the time t_{ij}^s when $W_{ij}(t, p) = \theta$. A process can have $0, 1, ..., s$ switching times.
The switching time set S_i for the ith variable contains all the switching times t_{ij}^s where $j = 1, .., k$.
The global switching time set S is the union of all S_i.

The relative weights evolve with time: for example a process can be *active* at the beginning of the system evolution and becomes *inactive* at the end. Fig. 1(A) shows the dynamics of the absolute values of the processes in (4) for $t \, \epsilon \, [0,24]$ (hours). In Fig. 1(B) is shown the dynamics of the absolute weights.

First step: eliminate the always inactive processes. We are looking for a function $g(x^r)$ which approximates the function $f(x)$ and with $N_P^g < N_P^f$, where N_P is the number of processes in the ODE system.

The threshold θ value must be chosen between the range $[0,1]$: a low threshold should be chosen to avoid neglecting too many processes and changing significantly the dynamics of the new system with respect to the original one. In our work a threshold $\theta = 0.1$ is set. After having applied this procedure for every kinetic equation, we obtain the reduced model where:
if $W_{ij}(x(t), p) < \theta \; \forall \; t \, \epsilon \, [0,T]$ then $f_{ij}(x(t), p) = 0$;
if not, $g_{ij}(x(t), p) = f_{ij}(x(t), p)$.
To test the quality of the reduced model $g(x^r)$, we numerically compute the global relative error between the original and the reduced models for each variable: it is defined for the ith variable as

$$\frac{\int |x_i(t) - x_i^r(t)| dt}{\int |x_i(t)| dt} \qquad (7)$$

where $x_i(t)$ and $x_i^r(t)$ are respectively the solutions of the original and reduced systems.
Taking (4) as an example, we see that the processes $f_{8,3}$, $f_{8,4}$, $f_{8,6}$ are always under θ and that they can be neglected. The processes $f_{8,1}$ and $f_{8,5}$ are always over θ, while $f_{8,2}$ crosses the threshold twice and its switching times ($t_{8,2}^1 = 0.2$ h and $t_{8,2}^2 = 19.8$ h) are collected in the set S_8.
Second step: visualize the dynamical activity of the remaining processes.
We qualitatively study the evolution of the intensity of the

A

d(T$_2$)/dt processes

Velocity (nM/h)

- second phosphorylation
- second de-phosphorylation
- formation of the complex
- splitting of the complex
- T$_2$ degradation
- linear T$_2$ degradation

Time (h)

B

Weights of d(T$_2$)/dt processes

Weight (-)

- second phosphorylation
- second de-phosphorylation
- formation of the complex
- splitting of the complex
- T$_2$ degradation
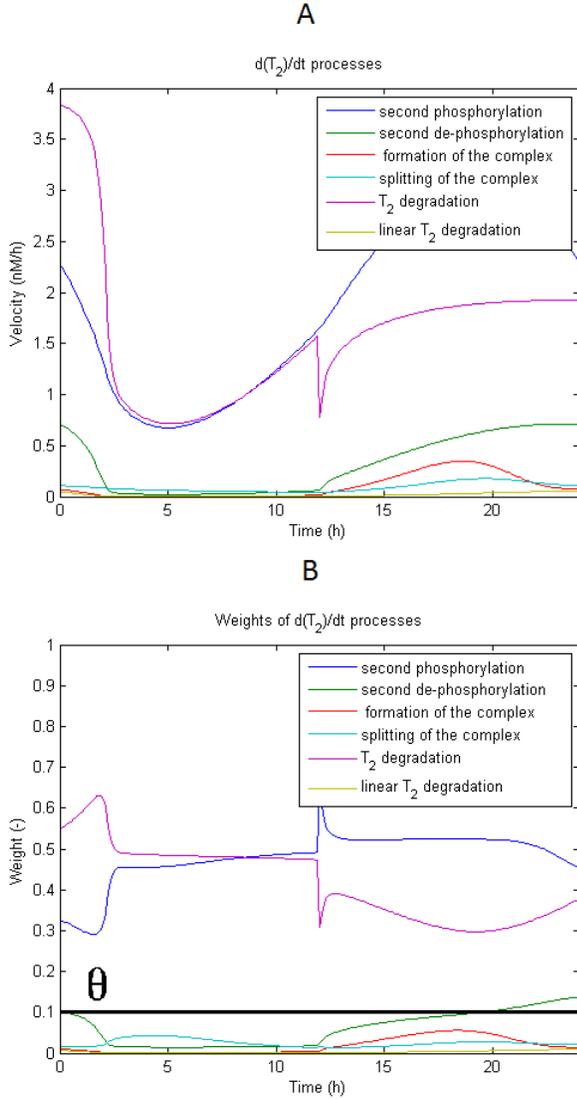- linear T$_2$ degradation

θ

Time (h)

Fig. 1. (A) Example of simulation of the processes of the 8th coordinate of the ODE system in [10] and (B) of the relative weights with the threshold θ set at 0.1. Three processes are *always inactive* during the time evolution of the system.

process *activity* in the reduced model $g(x^r)$. We use in this paper two graphic tools to summarize this information in one picture: a three dimensional bar graph and a heat map.
Third step: create a succession of sub-models.
The time period in which the system evolves can be split into time intervals using the switching times t_i previously grouped in the set S, sorted in ascending order: a succession of sub-models can be created, each one valid for a certain time window ($\Delta t_i = t_i - t_{i-1}$), which contains the core mechanisms in that period of time. To avoid having a large chain of sub-models, we reduce the number of time windows, by compacting with a clustering method the switching times that are close. We obtain a reduced number of time windows based on the reduced number of switching times: $\Delta t_i^r = t_i^r - t_{i-1}^r$ where $N_{\Delta t_i^r} << N_{\Delta t_i}$. A good

method to compact them in few clusters is the k-means clustering [12], which is really easy to compute in one dimension (time axis). Given our global switching time set (sorted in ascending order) $S = [t_1, t_2, ..., t_n]$, this method leads to group the n switching times into k ($\leq n$) clusters $C=\{C_1, C_2, ..., C_n\}$ so as to minimize the within-cluster sum of squares (WCSS):

$$\text{argmin}_C \sum_{i=1}^{k} \sum_{t \epsilon C_i} ||t - \mu_i||^2 \qquad (8)$$

where μ_i is the mean of the switching times in C_i. Following this operation, the cluster C_i is replaced by the mean μ_i, which is going to be the new switching time t_i^r, and we assume that all the processes that were going to switch their state in the cluster C_i are going to switch together at time t_i^r. There is no precise rule to choose k, but its value can be related with the difference between the maximum and minimum number of *active* processes during the time evolution of the system: if the difference is low, k should be chosen low as well. We finally end up with a small chain of sub-models that can be biologically interpreted.

III. MODEL FOR CIRCADIAN RHYTHMS IN DROSOPHILA

A. Description

We consider the model in [10] given by an ODE system of dimension 10. The model describes the circadian oscillations of the proteins timeless (TIM) and period (PER), which involve a negative feedback loop: the double phosphorylated forms of these proteins can be degraded in the cytoplasm or form the PER-TIM complex which, following its transport to the nucleus, inhibits the transcription of the *tim* and *per* genes and the subsequent accumulation of their mRNAs and proteins. The light sets the period of the oscillations to 24 hours precisely, by increasing the velocity of the degradation of the double phosphorylated form of TIM (from 2 nM/h to 4 nM/h in the model). Since the same oscillatory behavior is repeated every day, we focus our analysis on 24 hours only and use the model parameters given in Figures 2 and 4 of [10].

B. Model reduction

After having decomposed every kinetic equation of the model into a set of processes that we consider to be biologically relevant, we simulate their relative weight as shown before for (4) in Fig. 1. Every time that the weight of a process crosses the threshold θ, the corresponding process changes its state. The timing of the process *activation/inactivation* in each equation and the *active/inactive* processes along time is conveniently shown in a Temporal Process Map, as displayed in Fig. 2.
From this analysis, we obtain a reduced model by neglecting the processes that are *always inactive* (under the threshold $\theta = 0.1$ during the whole time). We hence neglect 18 of 44 processes: $f_{1,3}$, $f_{2,3}$, $f_{2,4}$, $f_{3,2}$, $f_{3,4}$, $f_{3,5}$, $f_{4,2}$, $f_{4,6}$, $f_{5,3}$, $f_{6,3}$, $f_{6,4}$, $f_{7,2}$, $f_{7,5}$, $f_{8,3}$, $f_{8,4}$, $f_{8,6}$, $f_{9,5}$, $f_{10,3}$. To test the quality of the reduced model, we calculate the global relative

error between the original and the reduced models for each variable using (7). The results in Table I show a good match between the original model and the reduced one.

C. Qualitative tool: Heat Map

In the previous analysis we used a Boolean approach to study the process *activity* or *inactivity* by means of the Temporal Process Map. We also have considered alternative graphical representations in this paper, since we also want to qualitatively understand, in one graph, the change over time of the intensity of the *activity/inactivity* of important biological processes. For the circadian model, we chose an Heat Map as a qualitative approach, where the individual values contained in a matrix are represented using colors. In the Heat Map that we generate, values along the rows are standardized, so that processes assume a red color (*active*) if their value is above the mean, black if their value is equal to the mean and green (*inactive*) if their value is below the mean of a column across all rows. If the red (or green) color is lighter, it means that the process is more *active* (or *inactive*). The Heat Map for the reduced model $g(x^r)$ is shown in Fig.3.
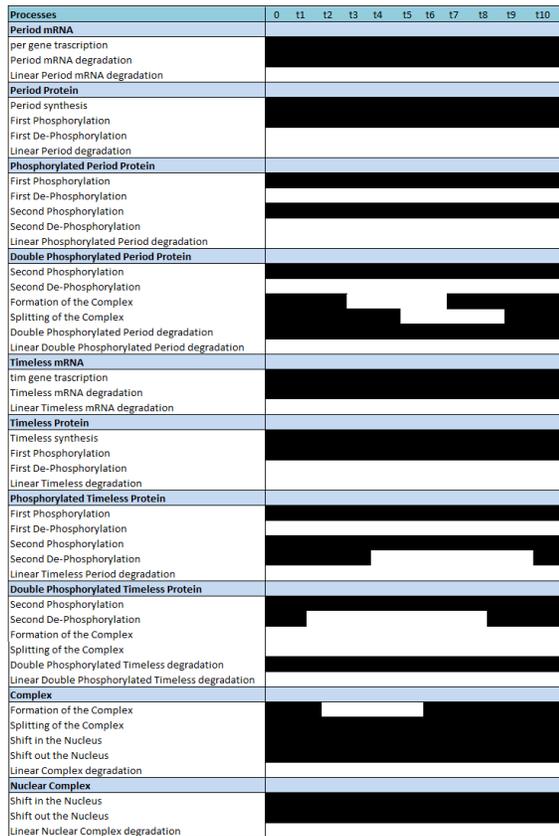
D. Creation of sub-models based on time windows

The Temporal Process Map is not only useful to identify negligible processes for model reduction, it can be used also to study the evolution of important (non-negligible) processes and create a succession of sub-models containing the core mechanisms, each one being valid for a certain time window.

Fig. 5(A) shows the *based-event* grid that is built, based on the switching times contained in the global switching time Set S, while Fig. 4 represents the number of *active* processes as a function of time for the Drosophila model. The minimum number of *active* processes is 21 and the maximum is 26 and

Fig. 2. Temporal Process Map of the process *activity* for the ODE model in Drosophila [10]. Black: *active* state. White: *inactive* state. Times t_i are, in the order, 0.2, 0.5, 1.3, 1.8, 6, 12.3, 15.8, 19.6, 19.8, and 21.5 h. The total duration of the simulation is 24 h.
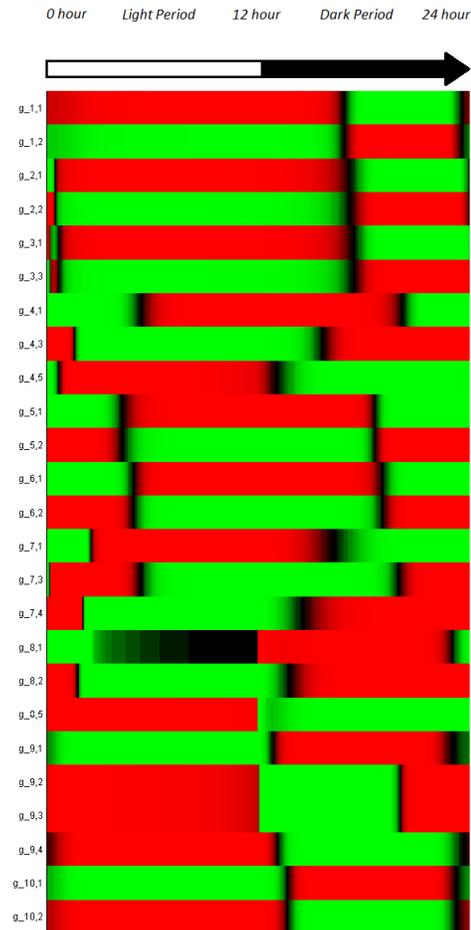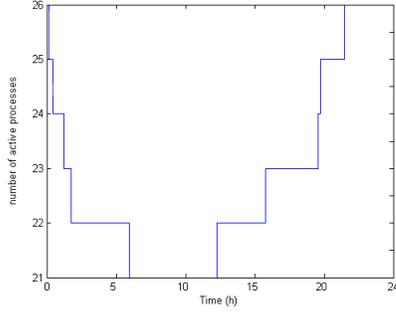

Fig. 3. Heat Map applied to the Drosophila Model.

Fig. 4. Evolution of the number of active processes as a function of time. The function increases or decreases at the switching time t_i in the time interval of the system evolution.
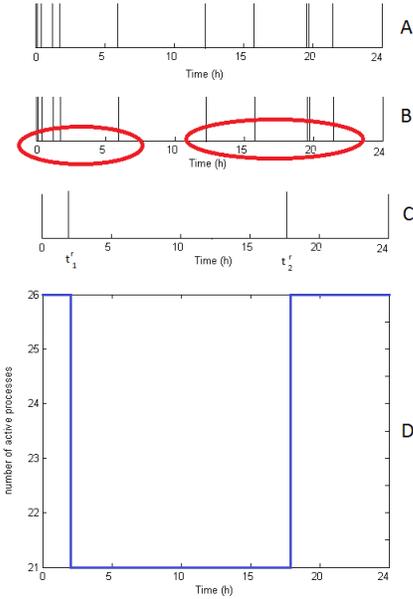

Fig. 5. (A) shows the *based-event* grid: every time window is delimited by the switching time t_i in the time interval of the system evolution (upper picture). In (B) the algorithm subsequently chooses the membership of every t_i in the Cluster C_1 or C_2 to minimize the WCCS expressed by (8) (middle picture). In (C) the cluster is then replaced with its centroid (in this case the mean) that will be the new approximate switching time. (D) shows the approximation of Fig. 4. The number of time windows becomes 2.
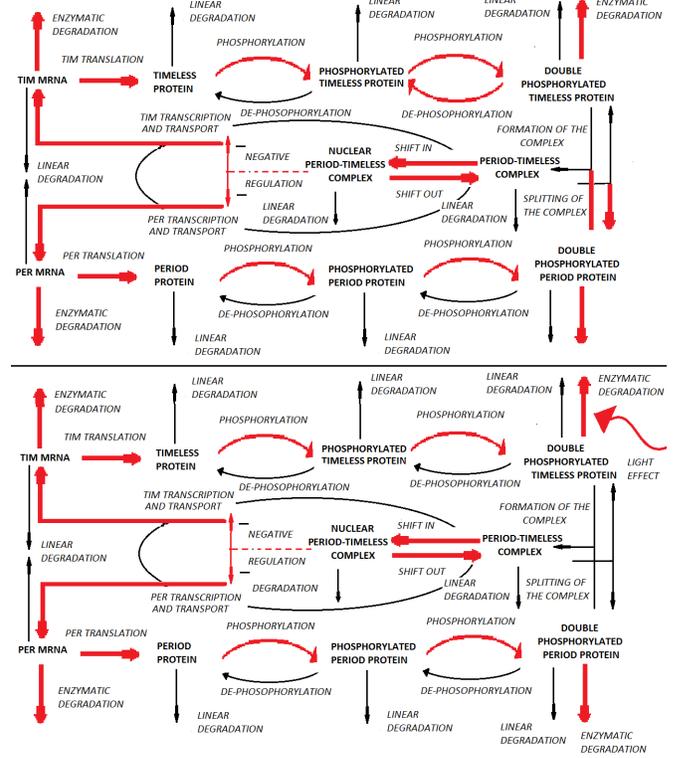

Fig. 6. Upper panel: Flux Map for the sub-model valid from 0 to 1.96 h and 17.8 h to 24 h. Lower panel: Flux Map for the sub-model valid from 1.96 h to 17.8 h. Black: *inactive* processes. Red: *active* processes.

good match between the original model and the first sub-model, and the original model and the second sub-model.

IV. MODEL FOR THE INFLUENCE OF RKIP ON THE ERK SIGNALING PATHWAY

A. Description

We consider the model in [11] of dimension 11. The ERK signaling pathway controls important cellular phenomena like proliferation or differentiation. The model describes the inhibition of the activation of RAF by RKIP, which regulates the ERK signaling pathway. We use in our analysis the parameters of [11] and initial values in [13].

Since no switches occur during the simulation (due to the strong *activity/inactivity* boundary separation and the absence

TABLE II

GLOBAL RELATIVE ERRORS FOR THE TWO SUB-MODELS OF THE DROSOPHILA MODEL

Variable	G. Rel. Err. S1 (%)	G. Rel. Err. S2 (%)
Period mRNA	13.63	7.70
Total Period Protein	1.61	7.06
Timeless mRNA	9.95	5.96
Total Timeless Protein	2.64	10.96
Complex	4.74	3.97
Nuclear Complex	5.36	5.85

the number of time windows is 10. We use the k-means clustering to compact together more switching times: we choose $k = 2$ because the difference between the maximum and minimum number of *active* processes is five (low value). The steps are shown in Fig. 5. We have thus created 2 sub-models: the first one (valid from 0 to 1.96 h and from 17.8 h to 24 h) coincides with the reduced model and the second one (valid from 1.96 h to 17.8 h) suppresses also the processes $f_{4,3}$, $f_{4,4}$, $f_{7,4}$, $f_{8,2}$, $f_{9,1}$.
In Fig. 6 is shown a Flux Map for the two sub-models, where red lines represent active processes in that time-window, while inactive ones are represented in black. The first sub-model essentially corresponds to a situation of night time and the second one, of day time. The results in Table II show a

of process weights crossing the threshold θ), we will show neither the Temporal Process Map nor the creation of sub-models.

B. Model reduction

We simulate the relative weight of the processes for every ODE as we did for the previous model. We are able to neglect 12 of 34 processes: $f_{1,2}$, $f_{2,2}$, $f_{3,2}$, $f_{3,4}$, $f_{4,2}$, $f_{5,3}$, $f_{6,3}$, $f_{7,2}$, $f_{8,2}$, $f_{9,2}$, $f_{10,2}$, $f_{11,2}$. The results in Table III show a really good match between the original model and the reduced one with a low global error. In addition to our study, other model reduction approaches with different goals have been applied to the ERK model in [11]: the quasi-steady-state-approximation used by Petrov *et al.* in [13] and the automatic complexity analysis by Lebiedz *et al.* in [14]. The work of [13] is concerned with the separation of variables with fast dynamics from those with slow dynamics with respect to a time scale and uses a mathematical scaling. The components of the ODEs related to the fast variables are expressed by algebraic equations. Analysing the resulting reduced model allowed the authors of the study to conclude that the variable m_4 (the complex Raf-1*/RKIP/ERK-PP) has the biggest influence on the system when it approaches its quasi-steady state. The work of [14] uses two different methods: the first one combines dynamic sensitivity analysis with singular value decomposition to find a minimal dimension of the model and the second one permits to reduce actually the dimension of the model and determine the variables which contribute more to the full dynamics of the system (variables m_5, m_8, m_{11}). Our analysis is different in the sense that we are not interested in which variable gives the bigger contribution to the dynamics of the full system but which mechanisms (processes) give the highest contribution to the dynamics of the variable: we study the influence of the processes. It is a reduction method that does not change in general the dimension and the biological structure of the system. For instance, the model in [11] is composed by two types of processes: *association processes*, where two or more proteins combine together to form a complex, and *dissociation processes* which correspond to the reverse mechanism. While the *association processes* are always *active* during the time evolution of the system, most of the *dissociation processes* can be considered negligible: it means that *association processes* play the bigger role in the dynamic for every variable of the system.

C. Qualitative tool:3-D bar Graph

We chose this time a three-dimensional bar graph as a qualitative approach to show the evolution of the intensity of the *activity* for every process (Fig.7). The height of the column gives the intensity of each process activity, while the color code indicates the weight of the process. In this graph we put the processes of the reduced model $g(x^r)$ during a one-second dynamics discretized in six time intervals. Fig.7 shows that the most *active* process during the system dynamics is the seventh process, the association of Raf-1*/RKIP and ERK-PP to form the Raf-1*/RKIP/ERK-PP

TABLE III
GLOBAL RELATIVE ERRORS FOR THE REDUCED ERK MODEL

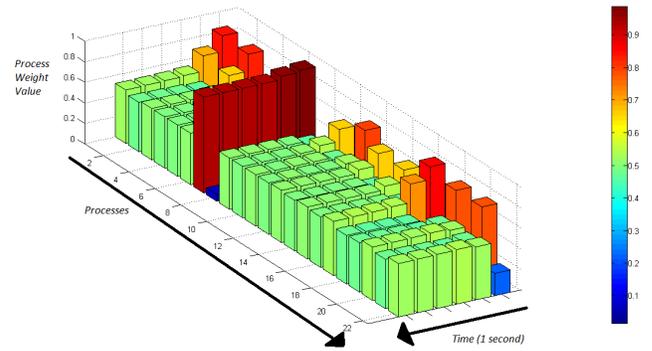| Variable	G. Rel. Err. (%)
Raf-1*	0.0012
RKIP	0.106
Raf-1*/RKIP	6.172
Raf-1*/RKIP/ERK-PP	0.0036
ERK-P	9.07
RKIP-P	0.138
MEK-PP	0.0056
MEK-PP/ERK	0.0148
ERK-PP	0.0013
RP	0.00049
RKIP-P	0.03451

Fig. 7. 3D bar graph: the x-axis represents the processes (the black arrow starts from the first process and ends with the last one), the y-axis corresponds to the time evolution and the z-axis, to the value of the weight.

complex. This process strongly influences the variable m_4, which was found by Petrov *et al.* to be the most important one for the system dynamics.

V. CONCLUSIONS

We have presented in this paper a method that allows to reduce biological models of high dimension and to analyze their core mechanisms and how these evolve dynamically, provided that their initial conditions and parameter values are known. Reduced models and sub-models obtained by this approach can be refined, by calibrating the reduced vector of parameters so as to obtain models that better approximate the original one. In a future step, Parameter Sensitivity Analysis will be applied to the reduced models and sub-models to test their robustness. Biological models of higher dimension with some parameter uncertainty (e.g. in some intervals) will be also analyzed in future works to study the consequences of this uncertainty on the dynamics of the sub-models.

APPENDIX

Full Drosophila model (see [10])

$$\frac{dM_P}{dt} = v_{sP}\frac{K_{IP}^n}{K_{IP}^n + C_N^n} - v_{mP}\frac{M_P}{K_{mP} + M_P} - k_d M_P$$

$$\frac{dP_0}{dt} = k_{sP}M_P - V_{1P}\frac{P_0}{K_{1P} + P_0} + V_{2P}\frac{P_1}{K_{2P} + P_1} - k_d P_0$$

TABLE IV
Variables of Drosophila model

Variable	Description
M_P	Period mRNA
P_0	Period Protein
P_1	Phosphorylated Period Protein
P_2	Double Phosphorylated Period Protein
M_T	Timeless mRNA
T_0	Timeless Protein
T_1	Phosphorylated Timeless Protein
T_2	Double Phosphorylated Timeless Protein
C	Complex
C_N	Nuclear Complex

TABLE V
Variables of ERK model

Variable	Description
m_1	Raf-1*
m_2	RKIP
m_3	Raf-1*/RKIP
m_4	Raf-1*/RKIP/ERK-PP
m_5	ERK-P
m_6	RKIP-P
m_7	MEK-PP
m_8	MEK-PP/ERK
m_9	ERK-PP
m_{10}	RP
m_{11}	RKIP-P

$$\frac{dP_1}{dt} = V_{1P}\frac{P_0}{K_{1P}+P_0} - V_{2P}\frac{P_1}{K_{2P}+P_1} - V_{3P}\frac{P_1}{K_{3P}+P_1} + V_{4P}\frac{P_2}{K_{4P}+P_2} - k_d P_1$$

$$\frac{dP_2}{dt} = V_{3P}\frac{P_1}{K_{3P}+P_1} - V_{4P}\frac{P_2}{K_{4P}+P_2} - k_3 P_2 T_2 + k_4 C - V_{dP}\frac{P_2}{K_{dP}+P_2} - k_d P_2$$

$$\frac{dM_T}{dt} = v_{sP}\frac{K_{IT}^n}{K_{IT}^n + C_N^n} - v_{mT}\frac{M_T}{K_{mT}+M_T} - k_d M_T$$

$$\frac{dT_0}{dt} = k_{sT}M_T - V_{1T}\frac{T_0}{K_{1T}+T_0} + V_{2T}\frac{T_1}{K_{2T}+T_1} - k_d T_0$$

$$\frac{dT_1}{dt} = V_{1T}\frac{T_0}{K_{1T}+T_0} - V_{2T}\frac{T_1}{K_{2T}+T_1} - V_{3T}\frac{T_1}{K_{3T}+T_1} + V_{4T}\frac{T_2}{K_{4T}+T_2} - k_d T_1$$

$$\frac{dT_2}{dt} = V_{3T}\frac{T_1}{K_{3T}+T_1} - V_{4T}\frac{T_2}{K_{4T}+T_2} - k_3 P_2 T_2 + k_4 C - V_{dT}\frac{T_2}{K_{dT}+T_2} - k_d T_2$$

$$\frac{dC}{dt} = k_3 P_2 T_2 - k_4 C - k_1 C + k_2 C_N - k_{dC} C$$

$$\frac{dC_N}{dt} = k_1 C - k_2 C_N - k_{dN} C_N$$

Reduced Drosophila model and first sub-model

$$\frac{dM_P}{dt} = v_{sP}\frac{K_{IP}^n}{K_{IP}^n + C_N^n} - v_{mP}\frac{M_P}{K_{mP}+M_P}$$

$$\frac{dP_0}{dt} = k_{sP}M_P - V_{1P}\frac{P_0}{K_{1P}+P_0}$$

$$\frac{dP_1}{dt} = V_{1P}\frac{P_0}{K_{1P}+P_0} - V_{3P}\frac{P_1}{K_{3P}+P_1}$$

$$\frac{dP_2}{dt} = V_{3P}\frac{P_1}{K_{3P}+P_1} - k_3 P_2 T_2 + k_4 C - V_{dP}\frac{P_2}{K_{dP}+P_2}$$

$$\frac{dM_T}{dt} = v_{sP}\frac{K_{IT}^n}{K_{IT}^n + C_N^n} - v_{mT}\frac{M_T}{K_{mT}+M_T}$$

$$\frac{dT_0}{dt} = k_{sT}M_T - V_{1T}\frac{T_0}{K_{1T}+T_0}$$

$$\frac{dT_1}{dt} = V_{1T}\frac{T_0}{K_{1T}+T_0} - V_{3T}\frac{T_1}{K_{3T}+T_1} + V_{4T}\frac{T_2}{K_{4T}+T_2}$$

$$\frac{dT_2}{dt} = V_{3T}\frac{T_1}{K_{3T}+T_1} - V_{4T}\frac{T_2}{K_{4T}+T_2} - v_{dT}\frac{T_2}{K_{dT}+T_2}$$

$$\frac{dC}{dt} = k_3 P_2 T_2 - k_4 C - k_1 C + k_2 C_N$$

$$\frac{dC_N}{dt} = k_1 C - k_2 C_N$$

Second Drosophila sub-model

$$\frac{dM_P}{dt} = v_{sP}\frac{K_{IP}^n}{K_{IP}^n + C_N^n} - v_{mP}\frac{M_P}{K_{mP}+M_P}$$

$$\frac{dP_0}{dt} = k_{sP}M_P - V_{1P}\frac{P_0}{K_{1P}+P_0}$$

$$\frac{dP_1}{dt} = V_{1P}\frac{P_0}{K_{1P}+P_0} - V_{3P}\frac{P_1}{K_{3P}+P_1}$$

$$\frac{dP_2}{dt} = V_{3P}\frac{P_1}{K_{3P}+P_1} - V_{dP}\frac{P_2}{K_{dP}+P_2}$$

$$\frac{dM_T}{dt} = v_{sP}\frac{K_{IT}^n}{K_{IT}^n + C_N^n} - v_{mT}\frac{M_T}{K_{mT}+M_T}$$

$$\frac{dT_0}{dt} = k_{sT}M_T - V_{1T}\frac{T_0}{K_{1T}+T_0}$$

$$\frac{dT_1}{dt} = V_{1T}\frac{T_0}{K_{1T}+T_0} - V_{3T}\frac{T_1}{K_{3T}+T_1}$$

$$\frac{dT_2}{dt} = V_{3T}\frac{T_1}{K_{3T}+T_1} - V_{dT}\frac{T_2}{K_{dT}+T_2}$$

$$\frac{dC}{dt} = -k_4 C - k_1 C + k_2 C_N$$

$$\frac{dC_N}{dt} = k_1 C - k_2 C_N$$

Full model ERK model (see [11])

$$\frac{dm_1}{dt} = -k_1 m_1 m_2 + k_2 m_3 + k_5 m_4$$

$$\frac{dm_2}{dt} = -k_1 m_1 m_2 + k_2 m_3 + k_{11} m_{11}$$

$$\frac{dm_3}{dt} = k_1 m_1 m_2 - k_2 m_3 - k_3 m_3 m_9 + k_4 m_4$$

$$\frac{dm_4}{dt} = k_3 m_3 m_9 - k_4 m_4 - k_5 m_4$$

$$\frac{dm_5}{dt} = k_5 m_4 - k_6 m_5 m_7 + k_7 m_8$$

$$\frac{dm_6}{dt} = k_5 m_4 - k_9 m_6 m_{10} + k_{10} m_{11}$$

$$\frac{dm_7}{dt} = -k_6 m_5 m_7 + k_7 m_8 + k_8 m_8$$

$$\frac{dm_8}{dt} = k_6 m_5 m_7 - k_7 m_8 - k_8 m_8$$

$$\frac{dm_9}{dt} = -k_3 m_3 m_9 + k_4 m_4 + k_8 m_8$$

$$\frac{dm_{10}}{dt} = -k_9 m_6 m_{10} + k_{10} m_{11} + k_{11} m_{11}$$

$$\frac{dm_{11}}{dt} = k_9 m_6 m_{10} - k_{10} m_{11} - k_{11} m_{11}$$

Reduced ERK model

$$\frac{dm_1}{dt} = -k_1 m_1 m_2 + k_5 m_4$$

$$\frac{dm_2}{dt} = -k_1 m_1 m_2 + k_{11} m_{11}$$

$$\frac{dm_3}{dt} = k_1 m_1 m_2 - k_3 m_3 m_9$$

$$\frac{dm_4}{dt} = k_3 m_3 m_9 - k_5 m_4$$

$$\frac{dm_5}{dt} = k_5 m_4 - k_6 m_5 m_7$$

$$\frac{dm_6}{dt} = k_5 m_4 - k_9 m_6 m_{10}$$

$$\frac{dm_7}{dt} = -k_6 m_5 m_7 + k_8 m_8$$

$$\frac{dm_8}{dt} = k_6 m_5 m_7 - k_8 m_8$$

$$\frac{dm_9}{dt} = -k_3 m_3 m_9 + k_8 m_8$$

$$\frac{dm_{10}}{dt} = -k_9 m_6 m_{10} + k_{11} m_{11}$$

$$\frac{dm_{11}}{dt} = k_9 m_6 m_{10} - k_{11} m_{11}$$

ACKNOWLEDGMENT

REFERENCES

[1] K. Bettenbrock, S. Fischer, A. Kremling, K. Jahreis, T. Sauter, and E.-D. Gilles, "A quantitative approach to catabolite repression in escherichia coli," *Journal of Biological Chemistry*, vol. 281, no. 5, pp. 2578–2584, 2006.

[2] L. Kuepfer, M. Peter, U. Sauer, and J. Stelling, "Ensemble modeling for analysis of cell signaling dynamics," *Nature biotechnology*, vol. 25, no. 9, pp. 1001–1006, 2007.

[3] D. Degenring, C. Froemel, G. Dikta, and R. Takors, "Sensitivity analysis for the reduction of complex metabolism models," *Journal of Process Control*, vol. 14, no. 7, pp. 729–745, 2004.

[4] L. A. Segel and M. Slemrod, "The quasi-steady-state assumption: a case study in perturbation," *SIAM review*, vol. 31, no. 3, pp. 446–477, 1989.

[5] H. K. Khalil, *Nonlinear systems*, 2nd ed. Prentice hall New Jersey, 1996.

[6] H. De Jong, J.-L. Gouzé, C. Hernandez, M. Page, T. Sari, and J. Geiselmann, "Qualitative simulation of genetic regulatory networks using piecewise-linear models," *Bulletin of mathematical biology*, vol. 66, no. 2, pp. 301–340, 2004.

[7] V. Baldazzi, D. Ropers, Y. Markowicz, D. Kahn, J. Geiselmann, and H. De Jong, "The carbon assimilation network in escherichia coli is densely connected and largely sign-determined by directions of metabolic fluxes," *PLoS computational biology*, vol. 6, no. 6, p. e1000812, 2010.

[8] L. Petzold and W. Zhu, "Model reduction for chemical kinetics: An optimization approach," *AIChE Journal*, vol. 45, no. 4, pp. 869–886, 1999.

[9] M. Apri, M. de Gee, and J. Molenaar, "Complexity reduction preserving dynamical behavior of biochemical networks," *Journal of theoretical biology*, vol. 304, pp. 16–26, 2012.

[10] J.-C. Leloup and A. Goldbeter, "A model for circadian rhythms in drosophila incorporating the formation of a complex between the per and tim proteins," *Journal of biological rhythms*, vol. 13, no. 1, pp. 70–87, 1998.

[11] C. Kwang-Hyun, S. Sung-Young, K. Hyun-Woo, O. Wolkenhauer, B. McFerran, and W. Kolch, "Mathematical modeling of the influence of rkip on the erk signaling pathway," in *Computational methods in systems biology*. Editors: C. Priami. Springer, 2003, pp. 127–141.

[12] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 881–892, 2002.

[13] V. Petrov, E. Nikolova, and O. Wolkenhauer, "Reduction of nonlinear dynamic systems with an application to signal transduction pathways," *Systems Biology, IET*, vol. 1, no. 1, pp. 2–9, 2007.

[14] D. Lebiedz, D. Skanda, and M. Fein, "Automatic complexity analysis and model reduction of nonlinear biochemical systems," in *Computational methods in systems biology*. Editors: M. Bernardo, P. Degano, G. Zavattano. Springer, 2008, pp. 123–140.