Brief paper

Symbolic reachability analysis of genetic regulatory networks using discrete abstractions[☆]

Grégory Batt[a,b,c,1], Hidde de Jong[a,*], Michel Page[a,e], Johannes Geiselmann[a,d]

[a]*INRIA Grenoble - Rhône-Alpes, 655 avenue de l'Europe, Montbonnot, 38334 Saint Ismier Cedex, France*
[b]*Center for Information and Systems Engineering, Boston University, USA*
[c]*Verimag, Grenoble, France*
[d]*Laboratoire Adaptation et Pathogénie des Microorganismes (CNRS UMR 5163), Université Joseph Fourier, Grenoble, France*
[e]*Ecole Supérieure des Affaires, Université Pierre Mendès France, Grenoble, France*

Abstract

We use hybrid-systems techniques for the analysis of reachability properties of a class of piecewise-affine (PA) differential equations that are particularly suitable for the modeling of genetic regulatory networks. More specifically, we introduce a hyperrectangular partition of the state space that forms the basis for a discrete abstraction preserving the sign of the derivatives of the state variables. The resulting discrete transition system provides a qualitative description of the network dynamics that is well-adapted to available experimental data and that can be efficiently computed in a symbolic manner from inequality constraints on the parameters.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Piecewise-affine differential equations; Qualitative analysis; Discrete abstraction; Hybrid systems; Genetic regulatory networks; Systems biology

1. Introduction

A class of piecewise-affine (PA) differential equations introduced by Glass and Kauffman (1973) in the seventies has been shown particularly suitable for modeling so-called *genetic regulatory networks*, networks of genes, proteins, small molecules, and their mutual interactions that are involved in the control of intracellular processes. The dynamics of these networks is hybrid in nature, in the sense that the continuous evolution of the concentration of proteins and other molecules is punctuated by discrete changes in the activity of genes coding for the proteins. This switch-like character of gene regulation is well-captured by the PA models, which have the additional advantage that the qualitative dynamics of the systems is relatively simple to

analyze, even in higher dimensions, without the use of numerical values for the kinetic parameters. Given that such information is usually absent in molecular biology, the PA models have been found to be a valuable tool for the practical analysis of complex genetic regulatory networks, which would be difficult to handle with more conventional nonlinear models.

The dynamical properties of the class of PA models considered here have been the subject of active research for more than three decades (e.g., Glass & Kauffman, 1973; Belta, Esposito, Kim, & Kumar, 2005; Edwards, 2000; Ghosh & Tomlin, 2004; Gouzé & Sari, 2002; Mestl, Plahte, & Omholt, 1995, see Batt, Ropers, de Jong, Page, & Geiselmann, 2007 for further references). In our previous work (de Jong et al., 2004), we have made a contribution to the analysis of these PA models by showing how to use differential inclusions to deal with discontinuities in the righthand side of the equations. Moreover, we have proposed algorithms and tools to compute a discrete representation of the state space dynamics in the form of a state transition graph.

In this note, we carry the analysis of the PA models further on a number of points, borrowing concepts and techniques from the field of hybrid systems. First, we partition the state space

into hyperrectangular regions in which the time derivatives of the solutions have a unique sign pattern. In a second step, this partition motivates the definition of a discrete abstraction (Alur, Henzinger, Lafferriere, & Pappas, 2000) that leads to a discrete transition system providing a finer-grained description of the qualitative dynamics of the system than was hitherto possible and which is better adapted to currently available experimental data. Third, we give rules for the symbolic computation of the discrete state transition system from inequality constraints on the parameters. The implementation of these rules has been shown to scale up to large and complex PA models of genetic regulatory networks.

A long version of this note, containing the proofs of all lemmas and propositions, as well as examples of the application of the method to an actual biological network, is available as supplemental material on the INRIA web site (Batt et al., 2007). This work extends a short and preliminary version of the paper presented at the HSCC conference (Batt et al., 2005).

2. PA systems

The dynamics of genetic regulatory networks can be described by PA differential equation models using step functions to account for regulatory interactions (Glass & Kauffman, 1973; Mestl et al., 1995). Fig. 1 gives an example of the PA model of a simple two-gene network. Below we define the models and review some mathematical properties.

We denote by $x = (x_1, \ldots, x_n)' \in \Omega$ a vector of cellular protein concentrations, where $\Omega = \Omega_1 \times \cdots \times \Omega_n \subset \mathbb{R}^n_{\geqslant 0}$ is a bounded n-dimensional hyperrectangular state space region. For each protein concentration x_i, $i \in \{1, \ldots, n\}$, we distinguish a set of constant, strictly positive threshold concentrations $\{\theta_i^1, \ldots, \theta_i^{p_i}\}$, $p_i > 0$. At its threshold concentrations a protein may affect the expression of genes encoding other proteins or the expression of its own gene, thus changing the regulatory mode of the system. The threshold concentrations induce a natural partition of Ω into hyperrectangular regions (de Jong et al., 2004).

Definition 1 (*Mode domain partition*). $\mathscr{M}$ is the hyperrectangular partition (Batt et al., 2007) of Ω induced by $\{\theta_i^1, \ldots, \theta_i^{p_i}\}$. The sets $M \in \mathscr{M}$ are called *mode domains*.

Fig. 2(a) shows the mode domain partition of the state space of the example network. We distinguish between mode domains like M^2 and M^7, which are located on (intersections of) threshold hyperplanes, and mode domains like M^1, which are not. The former are called *singular* and the latter *regular* mode domains. We denote by $\mathscr{M}_r$ and $\mathscr{M}_s$ the sets of regular and singular mode domains, respectively. Note that $\mathscr{M} = \mathscr{M}_r \cup \mathscr{M}_s$.

The PA models with step functions can be defined on the mode partition as follows (de Jong et al., 2004; Glass & Kauffman, 1973):

$$\dot{x} = h(x) = \mu^M - v^M x, \quad x \in M \in \mathscr{M}_r, \tag{1}$$

where μ^M is a vector of positive constants and $v^M = \mathrm{diag}(v_1^M, \ldots, v_n^M)$ a diagonal matrix of strictly positive constants. That is, in each mode domain the rate of change of
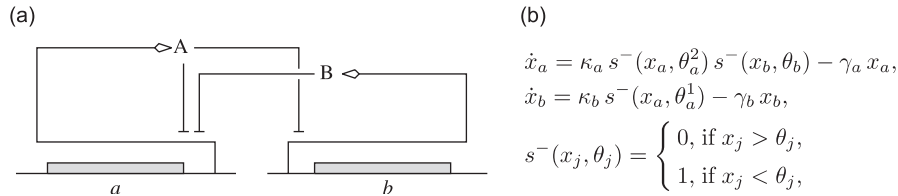

$$\dot{x}_a = \kappa_a \, s^-(x_a, \theta_a^2) \, s^-(x_b, \theta_b) - \gamma_a \, x_a,$$
$$\dot{x}_b = \kappa_b \, s^-(x_a, \theta_a^1) - \gamma_b \, x_b,$$
$$s^-(x_j, \theta_j) = \begin{cases} 0, \text{ if } x_j > \theta_j, \\ 1, \text{ if } x_j < \theta_j, \end{cases}$$

Fig. 1. (a) Example of a genetic regulatory network of two genes (*a* and *b*), each coding for a regulatory protein (A and B). Protein B inhibits the expression of gene *b*, while protein A inhibits the expression of gene *b* and its own gene. (b) PA model with step functions corresponding to the network in (a). Protein A is synthesized at a rate κ_a, if and only if the concentration of protein A is below its threshold θ_a^2 ($x_a < \theta_a^2$) and the concentration of protein B below its threshold θ_b ($x_b < \theta_b$). The degradation of protein A occurs at a rate proportional to the concentration of the protein itself ($\gamma_a x_a$).
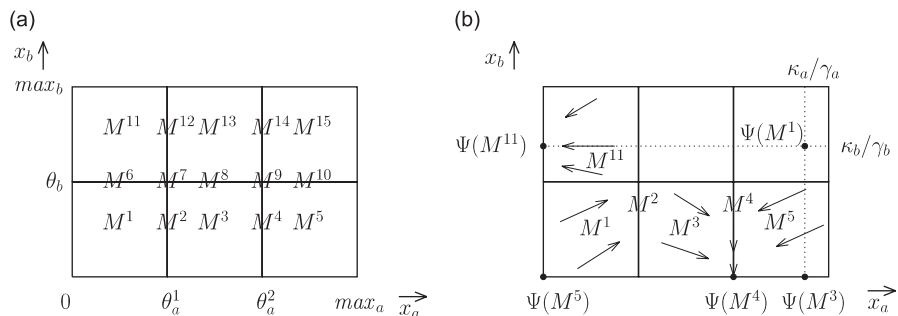

Fig. 2. (a) Mode domain partition of the state space for the model of Fig. 1(b). (b) Focal sets and vector fields associated with the mode domains $M^1, \ldots, M^5, M^{11}$. The use of differential inclusions gives rise to sliding mode solutions in M^4 (de Jong et al., 2004; Gouzé & Sari, 2002).

protein concentration x_i is defined as the difference of a synthesis rate μ^M and a degradation rate $v^M x$. For instance, by evaluating the step function expression in the model of Fig. 1(b) in mode domain M^1, we find that $\mu^{M^1} = (\kappa_a, \kappa_b)'$ and $v^{M^1} = \mathrm{diag}(\gamma_a, \gamma_b)$.

Because the mode of the system might change abruptly at threshold hyperplanes $\Theta = \bigcup_{i \in \{1,\ldots,n\}, l_i \in \{1,\ldots,p_i\}} \{x \in \Omega \mid x_i = \theta_i^{l_i}\}$ it is difficult to define the global dynamics of (1). To deal with this problem, we follow an approach similar to the one proposed by Filippov (1988). It consists in extending *differential equations* to *differential inclusions*. This requires the preliminary introduction of topological concepts. For every hyperrectangular set, $R \subseteq \Omega$, of dimension k, $k \in \{0, \ldots, n\}$, we define the supporting hyperplane of R, $supp(R) \subseteq \Omega$, as the k-dimensional hyperplane containing R. The *boundary* of R in $supp(R)$, denoted by ∂R, is defined as the set $\overline{R} / \overset{\circ}{R}$, where $\overline{R}$ is the closure of R and $\overset{\circ}{R}$ its interior. To every singular mode domain $M \in \mathcal{M}_s$, we associate the set $R(M)$ of the regular mode domains M' having M in their boundary, i.e., $R(M) = \{M' \in \mathcal{M}_r \mid M \subseteq \partial M'\}$.

We extend Eq. (1) to the differential inclusion

$$\dot{x} \in H(x), \quad x \in M \in \mathcal{M}, \tag{2}$$

$$H(x) = \begin{cases} \{\mu^M - v^M x\} & \text{if } x \in M \in \mathcal{M}_r, \\ \overline{rect}(\{\mu^{M'} - v^{M'} x \mid M' \in R(M)\}) & \text{otherwise,} \end{cases}$$

where $\overline{rect}(P)$ denotes the smallest closed hyperrectangular set containing the set P (de Jong et al., 2004).

We now define the *PA system* Σ as the triple (Ω, Θ, H). A *solution* of the PA system Σ on a time interval I is an absolutely continuous vector-valued function $\xi(t)$ such that $\dot{\xi}(t) \in H(\xi(t))$ almost everywhere on I. In particular, $\dot{\xi}(t) \in H(\xi(t))$ may not hold, if ξ reaches or leaves Θ at t. The existence of at least one solution ξ on some time interval $[0, \tau]$, $\tau > 0$, with initial condition $\xi(0) = x_0$ is guaranteed for all x_0 in Ω (Filippov, 1988). However, there is, in general, not a unique solution. The set Ξ_Σ is the set of all solutions $\xi(t)$ starting in Ω on a finite or infinite time interval of the PA system Σ. We restrict our analysis to the solutions in Ξ_Σ that reach and leave a threshold hyperplane finitely many times. The dynamics of Σ is thus formally defined by the set of solutions Ξ_Σ.

To every mode domain $M \in \mathcal{M}$ is associated a *focal set* $\Psi(M)$ (de Jong et al., 2004):

$$\Psi(M) = \begin{cases} \{\psi(M)\} & \text{if } M \in \mathcal{M}_r, \\ supp(M) \\ \quad \cap \overline{rect}(\{\psi(M') \mid M' \in R(M)\}) & \text{if } M \in \mathcal{M}_s, \end{cases}$$

where $\psi(M) = (v^M)^{-1} \mu^M$, $M \in \mathcal{M}_r$. Note that for singular mode domains $\Psi(M)$ is generally not a single point. Additionally, we will make the generic assumption that the focal sets $\Psi(M)$, for all $M \in \mathcal{M}_r$, are not located in the threshold hyperplanes Θ.

The reason for speaking of focal sets is that the solutions in a mode domain M monotonically converge towards (non-empty) focal sets $\Psi(M)$. This is formally stated in the following two lemmas.

Lemma 1 (*Glass & Kauffman, 1973*). *Let $M \in \mathcal{M}_r$. Every solution $\xi(t) \in \Xi_\Sigma$ in M monotonically converges towards $\Psi(M)$.*

Lemma 2. *Let $M \in \mathcal{M}_s$ and $\Psi(M) \neq \{\}$. For every solution $\xi(t) \in \Xi_\Sigma$ in M, and every $i \in \{1, \ldots, n\}$ such that M_i is included in a threshold hyperplane, it holds that $\dot{\xi}_i(t) = 0$. For all other i, $\xi_i(t)$ monotonically converges towards $\Psi_i(M)$, as long as $\xi_i(t) \notin \Psi_i(M)$.*

Finally, if $\Psi(M)$ is empty, then every solution passes through M instantaneously.

Lemma 3. *Let $M \in \mathcal{M}_s$. Every solution $\xi(t) \in \Xi_\Sigma$ arriving at M instantaneously crosses the domain if and only if $\Psi(M) = \{\}$.*

In the sequel, domains which every solution crosses instantaneously will be called *instantaneous*, whereas domains in which at least some solutions remain for some time will be called *persistent*. Fig. 2(b) shows the focal sets of four regular mode domains (M^1, M^3, M^5 and M^{11}) and two singular mode domains (M^2 and M^4).

The fact that every mode domain is associated with a unique focal set has provided the basis for a discrete abstraction criterion employed in our previous work (de Jong et al., 2004). However, this criterion disregards that the system does not always exhibit the same qualitative dynamics in different parts of a mode domain. Consider a solution $\xi(t) \in \Xi_\Sigma$ in M^{11} in Fig. 2(b): depending on whether $\xi_b(t)$ is larger than, equal to, or smaller than the focal concentration κ_b/γ_b, ξ_b will be decreasing, steady, or increasing. As a consequence, if we abstract the domain M^{11} away into a single discrete state, we will not be able to unambiguously infer that solutions entering this domain from M^6 are increasing in the x_b-dimension. In the context of model validation, this leads to over-conservative comparisons between model predictions and gene expression data provided by today's experimental techniques.

3. Flow domain partition

The mismatch between the mathematical analysis and the experimental data calls for a finer partitioning of the state space, which can then provide the basis for a more adequate abstraction criterion. Along these lines, the regular and singular mode domains distinguished above are repartitioned by means of the $(n - 1)$-dimensional hyperplanes corresponding to the focal concentrations.

Definition 2 (*Flow domain partition*). $\mathcal{D}^M$ is the hyperrectangular partition of a mode domain $M \in \mathcal{M}$ induced by $\{\psi_i(M)\}$ if M is regular, and by $\{\psi_i(M') \mid M' \in R(M)\}$ if M is singular, $i \in \{1, \ldots, n\}$. The *flow domain* partition of Ω is given by $\mathcal{D} = \bigcup_{M \in \mathcal{M}} \mathcal{D}^M$.
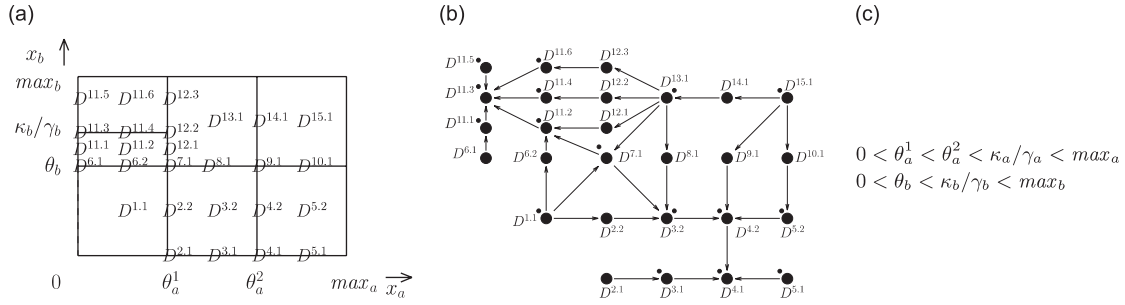
Fig. 3. (a) Flow domain partition of the state space of the model in Fig. 1(b), resulting in unique derivative sign patterns in each of the domains. For instance, for all $\xi(t)$ in $D^{11.2}$, it holds that $\dot{\xi}_a(t) < 0$ and $\dot{\xi}_b(t) > 0$ (compare with Fig. 2(b)). (b) State transition graph of the corresponding qualitative transition system. For the sake of clarity, self-transitions are represented by dots and transition labels are omitted. (c) Parameter inequality constraints for which the graph in (b) is invariant.

The partitioning of the state space into 27 flow domains is illustrated for the example system in Fig. 3(a). Every flow domain is included in a single mode domain, a relation captured by the surjective function $mode: \mathcal{D} \to \mathcal{M}$, where $mode(D) = M$, if and only if $D \subseteq M$. Similarly, the function $flow: \Omega \to \mathcal{D}$ denotes the surjective mapping that associates a point in the state space to its flow domain: $flow(x) = D$, if and only if $x \in D$.

The repartitioning of mode domain M^{11} leads to six flow domains (Fig. 3(a)). The finer partition guarantees that in every flow domain of M^{11}, the derivatives have a unique sign pattern. This property is true more generally. Consider a point x in a flow domain $D \in \mathcal{D}$. We denote by $S(x) \in 2^{\{-1,0,1\}^n}$ the set of derivative sign vectors of the solutions in D passing through x:

$$S(x) = \{\text{sign}(\dot{\xi}(\tau)) \mid \xi(t) \in \Xi_\Sigma \text{ in } D, \ \xi(\tau) = x,$$
$$\dot{\xi}(\tau) \in H(\xi(\tau))\}. \tag{3}$$

We find the following central result.

Theorem 1 (*Qualitatively identical dynamics*). *For all* $x, y \in D$ *and* $D \in \mathcal{D}$, $S(x) = S(y)$.

The main difficulty of the proof lies in the use of differential inclusions to deal with discontinuities on the threshold hyperplanes. Since the solutions of differential inclusions are not unique, several solutions may pass through x and their derivatives may have different signs in some dimension i. This situation does not occur in our two-gene example. We use the following lemma in the proof.

Lemma 4. *Let* $M \in \mathcal{M}_s$ *and* $\Psi(M) \neq \{\}$. *For every* $\psi \in \Psi(M)$ *and* $x \in M$, *there exists a solution* $\xi(t) \in \Xi_\Sigma$ *in* M, *with* $\xi(0) = x$, *such that* $\xi(t)$ *monotonically converges towards* ψ.

Proof of Theorem 1. Due to the hyperrectangular shape of $H(x)$, $S(x)$ can be decomposed into $S_1(x) \times \cdots \times S_n(x)$, where $S_i(x)$ denotes the set of the ith components of the derivative sign vectors of the solutions in D passing through x, $i \in \{1, \ldots, n\}$ (idem $S(y)$). Define $M = mode(D)$ and distinguish the cases (a)–(c).

(a) $M \in \mathcal{M}_r$. Suppose $1 \in S_i(x)$, but $1 \notin S_i(y)$, for any i (for 0 and −1 the argument is similar). By definition of $S_i(x)$ this

means that $\dot{\xi}_i(\tau) > 0$ for some $\xi(t) \in \Xi_\Sigma$ in D and $\tau \geqslant 0$, such that $\xi(\tau) = x$. On theother hand, $\dot{\varphi}_i(\sigma) \leqslant 0$ for all $\varphi(t) \in \Xi_\Sigma$ in D and $\sigma \geqslant 0$, such that $\varphi(\sigma) = y$. Given that $\dot{\xi}_i(\tau) \in H_i(\xi(\tau)) = \{v_i^M(\psi_i(M) - \xi_i(\tau))\}$, due to the definition of $S(x)$, it follows that $\xi_i(\tau) = x_i < \psi_i(M)$. Similarly, $\varphi_i(\sigma) = y_i \geqslant \psi_i(M)$. But then, by Definition 2, x and y are not located in the same flow domain D, contrary to the assertion of the theorem. Therefore, $1 \in S_i(x)$ implies $1 \in S_i(y)$. The converse is shown in the same way.

(b) $M \in \mathcal{M}_s$ and $\Psi(M) = \{\}$. No solution remains in D, so by definition $S(x) = S(y) = \{\}$.

(c) $M \in \mathcal{M}_s$ and $\Psi(M) \neq \{\}$. For every i such that D_i is located in a threshold or focal hyperplane, we must have $\dot{\xi}_i(t) = 0$ for solutions remaining in D. Consequently, $S_i(x) = \{0\}$ (idem $S_i(y) = \{0\}$). For all other i, suppose $1 \in S_i(x)$, but $1 \notin S_i(y)$ (the argument for 0 and −1 is similar). It follows that $\dot{\xi}_i(\tau) > 0$ for some $\xi(t) \in \Xi_\Sigma$ in D and $\tau \geqslant 0$, such that $\xi(\tau) = x$. However, $\dot{\varphi}_i(\sigma) \leqslant 0$ for all $\varphi(t) \in \Xi_\Sigma$ in D and $\sigma \geqslant 0$, such that $\varphi(\sigma) = y$. From $\dot{\xi}_i(\tau) \in H_i(\xi(\tau)) = \overline{rect}(\{v_i^{M'}(\psi_i(M') - \xi_i(\tau)) \mid M' \in R(M)\})$, we conclude that $\xi_i(\tau) = x_i < \psi_i(M')$, for some $M' \in R(M)$. Similarly, we must have $\dot{\varphi}_i(\sigma) \in H_i(\varphi(\sigma)) = \overline{rect}(\{v_i^{M'}(\psi_i(M') - \varphi_i(\sigma)) \mid M' \in R(M)\})$. Now, if there were some $M' \in R(M)$ such that $\varphi_i(\sigma) = y_i < \psi_i(M')$, then by Lemma 4 there would exist a solution such that $\dot{\varphi}_i(\sigma) > 0$. Since this contradicts the assumption, we conclude that $\varphi_i(\sigma) = y_i \geqslant \psi_i(M')$, for all $M' \in R(M)$. This implies by Definition 2, that x and y are not located in the same flow domain D. Therefore, $1 \in S_i(x)$ implies $1 \in S_i(y)$ (and conversely). □

Theorem 1 suggests that the partition of the state space introduced in this section can be used as an abstraction criterion better-adapted to the available experimental data on gene expression. This idea will be further developed in the next section.

4. Qualitative abstraction of the PA dynamics

4.1. Qualitative PA transition systems

As a preparatory step, we define a *continuous transition system* having the same reachability properties as the original PA

system Σ. Consider $x \in D$ and $x' \in D'$, where $D, D' \in \mathscr{D}$ are flow domains. If there exists a solution $\xi(t)$ of Σ passing through x at time $\tau \in \mathbb{R}_{\geqslant 0}$ and reaching x' at time $\tau' \in \mathbb{R}_{>0} \cup \{\infty\}$, without leaving $D \cup D'$ in the time interval $[\tau, \tau']$, then the absolute continuity of $\xi(t)$ implies that D and D' are either equal or contiguous. We consequently distinguish three types of continuous transitions, corresponding to the cases where $D = D'$, $D \subseteq \partial D'$, and $D' \subseteq \partial D$, respectively: *internal*, denoted by $x \xrightarrow{int} x'$, *dimension increasing*, denoted by $x \xrightarrow{dim^+} x'$, and *dimension decreasing*, denoted by $x \xrightarrow{dim^-} x'$. This leads to the following definition:

Definition 3 (*PA transition system*). Σ-TS $= (\Omega, L, \Pi, \rightarrow, \vDash)$ is the transition system corresponding to the PA system $\Sigma = (\Omega, \Theta, H)$, where:

(1) Ω is the state space;
(2) $L = \{int, dim^+, dim^-\}$ is a set of labels denoting the three different types of transitions;
(3) $\Pi = \{Dsign = S \mid S \in 2^{\{-1, 0, 1\}^n}\}$ is a set of propositions describing the signs of the derivatives (*Dsign*) of the concentration variables;
(4) $\rightarrow$ is the transition relation describing the continuous evolution of the system, defined by $\rightarrow \subseteq \Omega \times L \times \Omega$, such that $x \xrightarrow{l} x'$ if and only if there exists $\xi(t) \in \Xi_\Sigma$ and τ, τ', such that $0 \leqslant \tau < \tau'$, $\xi(\tau) = x$, $\xi(\tau') = x'$, and
  - if $l = int$, then for all $t \in [\tau, \tau']$: $\xi(t) \in flow(x) = flow(x')$,
  - if $l = dim^+$, then for all $t \in (\tau, \tau']$: $\xi(t) \in flow(x') \neq flow(x)$,
  - if $l = dim^-$, then for all $t \in [\tau, \tau')$: $\xi(t) \in flow(x) \neq flow(x')$;
(5) $\vDash$ is the satisfaction relation of the propositions in Π, defined by $\vDash \subseteq \Omega \times \Pi$, such that $x \vDash Dsign = S$ if and only if $S = S(x)$.

The satisfaction relation $\vDash$ thus associates to each point x in the state space a qualitative description of the dynamics of the system at x. It is straightforward to show that a PA system Σ and its corresponding PA transition system Σ-TS have equivalent reachability properties (Batt et al., 2007).

The continuous PA transition system has an infinite number of states and transitions, so that conventional tools for model checking cannot be used to verify system properties. However, we can define a discrete transition system, with a finite number of states and transitions, which preserves important properties of the qualitative dynamics of the system. In order to achieve this, we introduce the equivalence relation $\sim_\Omega \subseteq \Omega \times \Omega$ induced by the partition $\mathscr{D}$ of the state space: $x \sim_\Omega x'$ if and only if $flow(x) = flow(x')$. From Theorem 1 it follows that $\sim_\Omega$ is *proposition-preserving* (Alur et al., 2000), in the sense that for all $x, x' \in D$ and for all $\pi \in \Pi$, $x \vDash \pi$ if and only if $x' \vDash \pi$.

The discrete or *qualitative abstraction* of a PA transition system Σ-TS, called *qualitative PA transition system*, is now defined as the quotient transition system of Σ-TS, given the equivalence relation $\sim_\Omega$ (Alur et al., 2000). Denoting the qualitative

PA transition system by Σ-QTS, we have Σ-QTS $= (\Omega/_{\sim_\Omega}, L, \Pi, \rightarrow_{\sim_\Omega}, \vDash_{\sim_\Omega})$.

Proposition 1 (*Qualitative PA transition system*). *Let* Σ-QTS $= (\Omega/_{\sim_\Omega}, L, \Pi, \rightarrow_{\sim_\Omega}, \vDash_{\sim_\Omega})$ *be the qualitative PA transition system corresponding to the PA transition system* Σ-TS $= (\Omega, L, \Pi, \rightarrow, \vDash)$. *Then*

(1) $\Omega/_{\sim_\Omega} = \mathscr{D}$;
(2) $\rightarrow_{\sim_\Omega} \subseteq \mathscr{D} \times L \times \mathscr{D}$, *such that* $D \xrightarrow{l}_{\sim_\Omega} D'$ *if and only if there exists* $\xi(t) \in \Xi_\Sigma$ *and* τ, τ', $0 \leqslant \tau < \tau'$, *such that* $\xi(\tau) \in D$, $\xi(\tau') \in D'$, *and*
  - *if* $l = int$, *then for all* $t \in [\tau, \tau']$: $\xi(t) \in D = D'$,
  - *if* $l = dim^+$, *then for all* $t \in (\tau, \tau']$: $\xi(t) \in D' \neq D$,
  - *if* $l = dim^-$, *then for all* $t \in [\tau, \tau')$: $\xi(t) \in D \neq D'$;
(3) $\vDash_{\sim_\Omega} \subseteq \mathscr{D} \times \Pi$, *such that* $D \vDash Dsign = S$ *if and only if for all* $x \in D$, $S(x) = S$.

Notice that the transitions labeled by dim^+ or dim^- connect two different flow domains, since in Proposition 1 we require that $D \neq D'$. This corresponds to a continuous evolution of the system along which it switches from one flow domain to another. On the contrary, the transitions labeled by int correspond to the continuous evolution of the system in a single flow domain. The satisfaction relation $\vDash_{\sim_\Omega}$ associates to every flow domain a qualitative description of the derivatives of the solution.

Σ-QTS can be represented by a directed graph $G = (\mathscr{D}, \rightarrow_{\sim_\Omega})$, called the *state transition graph*. The paths in G represent the *runs* of the system (Batt et al., 2007). The state transition graph corresponding to the two-gene example is represented in Fig. 3(b). Notice that the problem mentioned at the end of Section 2, which illustrated the need for a refined partition, has been solved. Mode domain M^{11} is repartitioned into flow domains $D^{11.1}, \ldots, D^{11.6}$, and solutions entering $D^{11.1}$ or $D^{11.2}$ are necessarily increasing in the x_b-dimension.

It directly follows from the definitions of quotient transition system and simulation of transition systems (Alur et al., 2000) that Σ-QTS is a *simulation* of Σ-TS. The converse is not true in general, so that Σ-QTS and Σ-TS are not bisimilar (Batt et al., 2007). As a consequence, if there exists a run $(x^0, \ldots, x^m)$ of Σ-TS, then there also exists a run $(D^0, \ldots, D^m)$ of Σ-QTS such that $x^i \in D^i$, for all $i \in \{0, \ldots, m\}$. In other words, Σ-QTS is a *conservative approximation* of Σ-TS.

4.2. Invariance of qualitative PA transition systems in parameter space

In order to analyze the robustness of the qualitative dynamics, we introduce a second equivalence relation $\sim_\Gamma \subseteq \Gamma \times \Gamma$, defined on the parameter space $\Gamma \subseteq \mathbb{R}^q_{>0}$, with q the number of parameters in Σ. Two parameter vectors $p, p' \in \Gamma$ are equivalent, if their corresponding qualitative PA transition systems, and hence their state transition graphs, are isomorphic. We denote by $\Gamma/_{\sim_\Gamma}$ the quotient parameter space. That is, $\Gamma/_{\sim_\Gamma}$ is a partition of the parameter space consisting of sets over which the qualitative PA transition system is invariant.

In particular, subsets of Γ defined by the following *parameter inequality constraints* are interesting, since they provide an invariance criterion for Σ-QTS.

Definition 4 (*Parameter inequality constraints*). The parameter inequality constraints of Σ are a set of total strict orders on $\{\theta_i^1, \ldots, \theta_i^{p_i}\} \cup \{\psi_i(M) \mid M \in \mathscr{M}_r\}$, for every $i \in \{1, \ldots, n\}$.

Proposition 2 (*Invariance*). *Let $P \subseteq \Gamma$ be a set defined by the parameter inequality constraints of Σ. Then, there exists some $Q \in \Gamma/_{\sim_\Gamma}$ such that $P \subseteq Q$.*

For the inequality constraints in Fig. 3(c), the state transition graph in Fig. 3(b) is invariant. This is important in practice, as numerical parameter values are usually not available. The weaker information required for the ordering of the threshold and focal parameters can often be obtained from the experimental literature.

5. Symbolic computation of qualitative PA transition system

The computation of the qualitative PA transition system Σ-QTS is greatly simplified by the fact that the domains D and the focal sets $\Psi(M)$ are hyperrectangular, which allows them to be expressed as product sets, i.e., $D = D_1 \times \cdots \times D_n$ and $\Psi(M) = \Psi_1(M) \times \cdots \times \Psi_n(M)$. As a consequence, the computation can be carried out for each dimension separately. In this section, we will describe rules to determine the set of states $\mathscr{D}$, the satisfaction relation $\vDash_{\sim_\Omega}$, and the transition relation $\rightarrow_{\sim_\Omega}$.

5.1. Computation of states

In order to compute the states of Σ-QTS, we need to determine the flow domain partition $\mathscr{D}$ of Ω (Proposition 1). This requires a total ordering of the threshold concentrations $\{\theta_i^1, \ldots, \theta_i^{p_i}\}$ and the focal concentrations $\{\psi_i(M) \mid M \in \mathscr{M}_r\}$, $i \in \{1, \ldots, n\}$ (Definition 2). The parameter inequality constraints of Σ provide this information.

The following two propositions, which are direct consequences of the mathematical properties of solutions of PA systems (Section 2), show how to compute the derivative sign patterns associated to each flow domain.

Proposition 3 (*Computation of Dsign*). *Let $D \in \mathscr{D}$ be an instantaneous flow domain. $D \vDash_{\sim_\Omega} Dsign = \{\}$.*

Proposition 4 (*Computation of Dsign*). *Let $D \in \mathscr{D}$ be a persistent flow domain. $D \vDash_{\sim_\Omega} Dsign = S \neq \{\}$, $S = S_1 \times \cdots \times S_n$, and for every $i \in \{1, \ldots, n\}$,*

    if D_i is located in a threshold or focal concentration hyperplane, then $S_i = \{0\}$;

    otherwise,

- *$-1 \in S_i$, if and only if for all $x \in D$ there exists $\psi \in \Psi(mode(D))$ such that $\psi_i - x_i < 0$;*

- *$0 \in S_i$, if and only if for all $x \in D$ there exists $\psi \in \Psi(mode(D))$ such that $\psi_i - x_i = 0$;*
- *$1 \in S_i$, if and only if for all $x \in D$ there exists $\psi \in \Psi(mode(D))$ such that $\psi_i - x_i > 0$.*

Notice that the inequality constraints of Definition 4 allow one to symbolically compute $\Psi(mode(D))$ and decide which of the conditions in the second part of Proposition 4 are satisfied. We illustrate the application of Propositions 3 and 4 to our example in Fig. 4.

5.2. Computation of transitions between states

In Section 4.1 we have distinguished three types of transitions: int, dim^-, and dim^+. We will formulate, for each of these three cases, transition rules that can be applied by means of the parameter inequality constraints of Σ.

The int transitions are easy to determine, since by Proposition 1 they are necessarily self-transitions that occur if and only if D is persistent.

Proposition 5 (*Computation of int transition*). *Let $D \in \mathscr{D}$. $D \xrightarrow{int}_{\sim_\Omega} D$ if and only if D is persistent.*

Recall that the persistence of a domain can be determined by checking whether $\Psi(mode(D)) \neq \{\}$ (Lemmas 1 and 3). Fig. 5 gives an example.

A dim^+ transition $D \xrightarrow{dim^+}_{\sim_\Omega} D'$ is dimension increasing, that is, $D \subseteq \partial D'$ (Section 4.1). In order to make D' reachable from D, the solutions in D' must point away from D in the dimensions $i \in \{1, \ldots, n\}$ for which $D_i \subseteq \partial D_i'$. This is expressed by the following proposition (the rule for dim^- transitions is almost symmetric, see Batt et al., 2007).

Proposition 6 (*Computation of dim^+ transition*). *Let $D, D' \in \mathscr{D}$ and $D \subseteq \partial D'$. $D \xrightarrow{dim^+}_{\sim_\Omega} D'$ if and only if $\Psi(mode(D')) \neq \{\}$ and there exist $x \in D$, $x' \in D'$, and $\psi' \in \Psi(mode(D'))$, such that for all $i \in \{1, \ldots, n\}$ for which $D_i \subseteq \partial D_i'$, it holds that $(\psi_i' - x_i)(x_i' - x_i) > 0$.*

Proof. Let $M' = mode(D')$. We first prove necessity by contraposition. If $\Psi(M') = \{\}$, then no solutions remain in D' and a transition from D to D' is not possible. Otherwise, suppose that for all $x \in D$, $x' \in D'$, and $\psi' \in \Psi(M')$, there exists some $i \in \{1, \ldots, n\}$ for which $D_i \subseteq \partial D_i'$ and $(\psi_i' - x_i)(x_i' - x_i) \leqslant 0$. Furthermore, assume $x_i' - x_i > 0$ for all $x \in D$, $x' \in D'$ (the case $x_i' - x_i < 0$ goes analogously). As a consequence, $\psi_i' \leqslant x_i < x_i'$, for all $x \in D$, $x' \in D'$, and $\psi' \in \Psi(M')$. By Lemmas 1 and 2, for all solutions $\xi(t) \in \Xi_\Sigma$ in D', $\xi_i(t)$ monotonically converges towards $\Psi_i(M')$. Therefore, no solution enters D' from D and there does not exist a transition $D \xrightarrow{dim^+}_{\sim_\Omega} D'$.

Next, we prove sufficiency. Suppose $\Psi(M') \neq \{\}$ and there exist $x \in D$, $x' \in D'$, and $\psi' \in \Psi(M')$, such that for all $i \in \{1, \ldots, n\}$ for which $D_i \subseteq \partial D_i'$, it holds that $(\psi_i' - x_i)(x_i' - x_i) > 0$. By the definition of the flow partition (Definition 2)
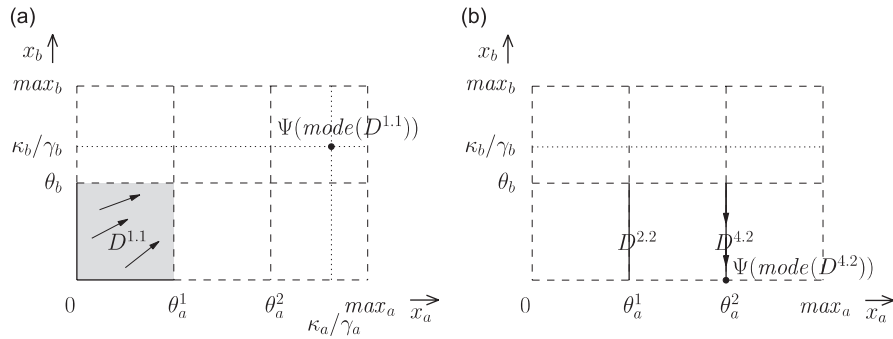
Fig. 4. Derivative sign patterns of flow domains. (a) Domain $D^{1.1}$. From the parameter inequality constraints in Fig. 3(c), it follows that $\psi_a - x_a > 0$ and $\psi_b - x_b > 0$, for all $x \in D^{1.1}$ and $\psi \in \Psi(mode(D^{1.1})) = \{(\kappa_a/\gamma_a, \kappa_b/\gamma_b)'\}$. Application of Proposition 4 then yields $Dsign(D^{1.1}) = \{(1, 1)'\}$. (b) Domains $D^{4.2}$ and $D^{2.2}$. $D^{2.2}$ is instantaneous, so from Proposition 3 it follows that $Dsign(D^{2.2}) = \{\}$. $D^{4.2}$ is located on the threshold $x_a = \theta_a^2$, while $\psi_b - x_b < 0$, for all $x \in D^{4.2}$ and $\psi \in \Psi(mode(D^{4.2})) = \{(\theta_a^2, \kappa_b/\gamma_b)'\}$. As a consequence, according to Proposition 4, $Dsign(D^{4.2}) = \{(0, -1)'\}$.

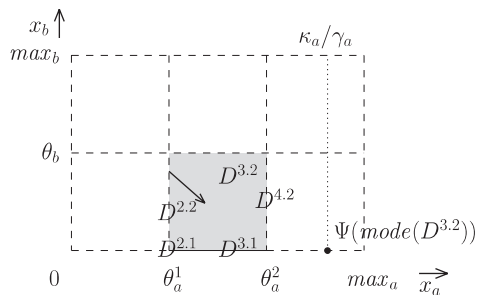Fig. 5. *int* and *dim*$^+$ transitions. $D^{3.2}$ is a persistent state for $\Psi(mode(D^{3.2})) \neq \{\}$. Therefore, $D^{3.2} \xrightarrow{int} _{\sim_\Omega} D^{3.2}$ (Proposition 5). Since $\psi_a' - x_a' > 0$ and $x_a' - x_a > 0$, for all $x \in D^{2.2}$, $x' \in D^{3.2}$, and $\psi' \in \Psi(mode(D^{3.2}))$, we have the transition $D^{2.2} \xrightarrow{dim^+} _{\sim_\Omega} D^{3.2}$ by Proposition 6.

this holds for all $x \in D$ and $x' \in D'$. We further assume that $x_i' - x_i > 0$ for all $x \in D$, $x' \in D'$ (the case $x_i' - x_i < 0$ goes analogously). As a consequence, $x_i < x_i' < \psi_i'$ for all $x \in D$ and $x' \in D'$, and some $\psi' \in \Psi(M')$. By Lemma 4, for all $x' \in D'$ there exists a solution $\xi(t) \in \Xi_\Sigma$ in D', with $\xi(0) = x'$, which monotonically converges towards ψ'. Thus, for all $x' \in D'$ for which $\xi_i(t) = x_i'$ and $x_i' < \psi_i'$, we have $\dot{\xi}_i(t) > 0$ and solutions can enter D' from D. As a consequence, there exists a transition $D \xrightarrow{dim^+} _{\sim_\Omega} D'$. $\square$

The total strict ordering defined by the parameter inequality constraints allows the necessary and sufficient conditions for the existence of a *dim*$^+$ transition to be straightforwardly tested. Fig. 5 gives an illustration of a *dim*$^+$ transition.

In summary, given a PA system Σ and parameter inequality constraints, we can compute the qualitative PA transition system Σ-QTS by means of Propositions 3–6. Instead of numerically computing the derivative signs in the domains and the transitions between domains, our symbolic algorithms reduce to simple checks of ordering relations, exploiting the parameter inequality constraints of Definition 4. The computation of Σ-QTS has been implemented in a new version of the

tool *Genetic Network Analyzer* (*GNA*). To facilitate the analysis of Σ-QTS, the state transition graph generated by GNA can be exported to standard model-checking tools (Batt, Ropers, de Jong, Geiselmann, Mateescu et al., 2005).

6. Discussion

We have presented a method for the qualitative analysis of a class of PA models that has been well-studied in mathematical biology. By defining a qualitative abstraction preserving the sign pattern of the derivatives of the state variables, the PA model is transformed into a discrete transition system whose properties can be analyzed by means of classical model-checking tools. The discrete transition system provides a conservative approximation of the qualitative dynamics of the system and can be easily computed in a symbolic manner by exploiting inequality constraints on the parameters.

The results of this paper extend our previous work on the qualitative analysis of PA models of genetic regulatory networks (de Jong et al., 2004). In particular, we have defined a fine-grained partition of the state space which underlies a discrete abstraction preserving stronger properties of the qualitative dynamics of the system, i.e., the derivative sign pattern. This has given rise to the definition and computation of a discrete transition system better adapted to the kind of experimental data available, in the sense that it avoids over-conservative verification of dynamical properties. Judging from our experience with several PA models of bacterial regulatory networks, the use of a finer-grained abstraction leads to only a modest increase in the size of the state transition graph (Batt et al., 2007). That is, the increase in precision does not exclude the application of the refined abstraction to larger systems.

The PA models in this paper have formed the basis for other work in the field of hybrid systems (e.g., Ghosh & Tomlin, 2004). Contrary to Ghosh and Tomlin (2004), we take into account the dynamics of the system on threshold hyperplanes, where equilibrium points and other phenomena of interest may occur (Gouzé & Sari, 2002). In Ghosh and Tomlin (2004) the partition of the state space is dynamically refined, so as to

arrive at a discrete transition system that is a better approximation of the original PA system. This requires the use of quantifier elimination methods which allow to decide more general and more powerful properties than the rules proposed in Section 5 of this paper, but that also incur higher computational costs and limit the scalability of the approach. Other groups have employed piecewise-multi affine models (Belta et al., 2005). These are more difficult to analyze and in higher dimensions usually require the application of numerical techniques. This is not straightforward to achieve for most biological systems, since numerical information on parameter values is usually imprecise or simply not available.

Acknowledgements

References

Alur, R., Henzinger, T. A., Lafferriere, G., & Pappas, G. J. (2000). Discrete abstractions of hybrid systems. *Proceedings of the IEEE*, *88*, 971–984.

Batt, G., Ropers, D., de Jong, H., Geiselmann, J., Mateescu, R., Page, M. et al. (2005). Validation of qualitative models of genetic regulatory networks by model checking. *Bioinformatics*, *21*, i19–i28.

Batt, G., Ropers, D., de Jong, H., Geiselmann, J., Page, M., & Schneider, D. (2005). Qualitative analysis and verification of hybrid models of genetic regulatory networks. In M. Morari, & L. Thiele (Eds.), *Proceedings of HSCC 2005. Lecture notes in computer science* (Vol. 3414, pp. 134–150). Berlin: Springer.

Batt, G., Ropers, D., de Jong, H., Page, M., & Geiselmann, J. (2007). *Symbolic reachability analysis of genetic regulatory networks using qualitative abstractions*. Technical Report RR-6136, INRIA, ⟨http://hal.inria.fr⟩.

Belta, C., Esposito, J. M., Kim, J., & Kumar, V. (2005). Computational techniques for analysis of genetic network dynamics. *International Journal of Robotics Research*, *24*, 219–235.

de Jong, H., Gouzé, J.-L., Hernandez, C., Page, M., Sari, T., & Geiselmann, J. (2004). Qualitative simulation of genetic regulatory networks using piecewise-linear models. *Bulletin of Mathematical Biology*, *66*, 301–340.

Edwards, R. (2000). Analysis of continuous-time switching networks. *Physica D*, *146*, 165–199.

Filippov, A. F. (1988). *Differential equations with discontinuous righthand sides*. Dordrecht: Kluwer Academic Publishers.

Ghosh, R., & Tomlin, C. J. (2004). Symbolic reachable set computation of piecewise affine hybrid automata and its application to biological modelling. *Systems Biology*, *1*, 170–183.

Glass, L., & Kauffman, S. A. (1973). The logical analysis of continuous non-linear biochemical control networks. *Journal of Theoretical Biology*, *39*, 103–129.

Gouzé, J.-L., & Sari, T. (2002). A class of piecewise linear differential equations arising in biological models. *Dynamical Systems*, *17*, 299–316.

Mestl, T., Plahte, E., & Omholt, S. W. (1995). A mathematical framework for describing and analysing gene regulatory networks. *Journal of Theoretical Biology*, *176*, 291–300.
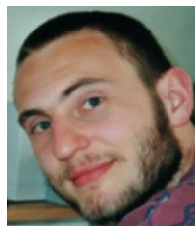
Grégory Batt obtained a BSc in Molecular and Cellular Biology from Ecole Normale Supérieure in Lyon and an MSc in Computer Science from the Joseph Fourier University (UJF) in Grenoble (France). After his PhD in Computer Science at UJF and INRIA Grenoble-Rhône-Alpes, obtained in 2005, he has continued his research on the interface of hybrid systems theory and systems biology with post-doctoral fellowships at Boston University and VERIMAG (CNRS, UJF). He is currently research scientist at INRIA/Paris (Rocquencourt)

Hidde de Jong obtained MSc degrees in Computer Science, Philosophy of Science, and Management Science from the University of Twente (the Netherlands) and completed a PhD thesis in Computer Science at the same university. He joined INRIA in 1998 and is currently a senior research scientist in the bioinformatics and biological modeling group at INRIA Grenoble-Rhône-Alpes (Grenoble).

Michel Page obtained a PhD thesis in Computer Science from the Institut National Polytechnique in Grenoble (INPG) in 1991. He is currently an associate professor in Computer Science at the Pierre Mendès-France University (Grenoble) and also affiliated with INRIA Grenoble-Rhône-Alpes, where he works on the development of modeling and simulation tools for bioinformatics.

Hans Geiselmann obtained a PhD in Molecular Biology from the University of Oregon (USA) in 1989. He has been an assistant professor at the University of Geneva and, since 1998, a full professor in Microbiology at UJF, where he explores interdisciplinary approaches to study the molecular mechanisms underlying the control of bacterial gene expression.