Embedding of biological regulatory networks and property preservation

Mbarka Mabrouki, Marc Aiguier, Jean-Paul Comet, Pascale Le Gall, Adrien Richard

Embedding of biological regulatory networks and property preservation

Mbarka Mabrouki, Marc Aiguier, Jean-Paul Comet,
Pascale Le Gall and Adrien Richard

Abstract. In the course of understanding biological regulatory networks (BRN), scientists usually start by studying small BRNs that they believe to be of particular importance to represent a biological function, and then, embed them in a larger network. Such a reduction can lead to neglect relevant regulations and to study a network whose properties can be very different from the properties of this network viewed as a part of the whole. In this paper we study, from a logical point of view, on which conditions concerning both networks, properties can be inherited by BRNs from sub-BRNs. We give some conditions on the nature of the network embeddings ensuring that dynamic properties on the embedded sub-BRNs are preserved at the level of the whole BRN.

Keywords. Biological regulatory networks, network embedding, property preservation, mathematical modeling, temporal logic.

1. Introduction

Framework

Systems biology is the research field that aims at understanding biological complex systems. It does not investigate individual genes, proteins or RNA at a time, but it studies the behaviors and relationships of all these components in a cell. The collection of these relationships is called Biological Regulatory Network (BRN). Among these BRNs, genetic regulatory networks are of particular interest because they react to the environment changes and regulate, for instance, the growth, replication and death of cells.

The complexity of the interactions between constituents of BRNs (mainly genes and proteins) makes intuitive reasoning difficult [28, 10, 26]. Modeling frameworks and simulation techniques are often used to analyze and understand such complex biological regulatory networks. Nevertheless, simulation techniques are not well suited for managing biological systems, which are large, complex and partially known. Indeed, the lack of precise knowledge about the system (what constituents/interactions are taken into account? What values are given to parameters? What is the confidence on these parameters? etc.) constitutes a major difficulty to handle computationally all possible hypotheses on the system.

Jacob and Monod [21] showed that cell behaviors are governed by their regulatory genes, which can be regarded as switches turning on or off the involved genes. Then, analysis of their behaviors can be done using qualitative modeling frameworks [12, 19, 29, 30, 17]. These qualitative frameworks consist in abstracting continuous concentrations of constituents into qualitative ones (discrete and finite) preserving qualitative observations (like presence/absence of a constituent, increasing of the concentration of a target when increasing the one of a regulator...). Moreover, the qualitative feature of such frameworks is well adapted to represent qualitative reasoning which characterizes a large fragment of biological studies. In fact, this emphasizes the logical relationships between the network elements, that can be incorporated into a logical framework to facilitate the study of biological systems [5, 4].

In this paper we consider the multi-valued discrete approach developed by R. Thomas and co-workers [30], where constituents concentrations are abstracted by integers denoting the thresholds which trigger interactions with other components in the network. In this formalism, biological systems are described by an interaction graph defining the static part of the system. A huge but finite set of state transition graphs can be built defining all the possible evolutions of the system. However, given an interaction graph, there are few dynamic models meeting the set of biological experiment observations and bringing into play interactions between constituents. In order to cut down in the class of dynamic models and to preserve only the good candidates, some recent works expressed these biological experiment observations by temporal properties [5, 4] and they used various model-checking techniques to select suitable dynamic models [14, 11, 5, 4, 25, 15, 3, 8, 22, 24].

Motivation

These logical approaches based on model-checking techniques have been shown very efficient to study small BRNs but are not well-adapted for large BRNs. Indeed, we have to deal with the classical state explosion problem: when the number of considered components is growing, the number of both models and states in each model are growing exponentially. This intrinsic limitation explains why in practice, only small BRNs are considered. This raises the following essential issue: what are the results stated on small BRNs which remain valid on larger ones containing them? More precisely, what are the conditions to impose on the embedding of a sub-BRN into a larger one in order to ensure that properties expressed on the

considered sub-BRN are preserved at the level of the larger BRN? Unfortunately, biological regulatory networks are very complex systems where combinations of sub-BRNs often lead to a global system getting some behaviors that cannot be defined as direct combinations of behaviors of sub-BRNs. The emergence of such behaviors results from the way sub-BRNs are grouped together. Thus, without knowing if there exist or not such emergent properties attached to a global BRN, questioning temporal properties (i.e. biological experimental observations) leads us to study the dynamics of the global BRN "from scratch", i.e without taking benefit of the dynamics of the sub-BRNs, which can be unacceptable at runtime. On the contrary, if we can state that in some particular cases, there is no emergent property when combining sub-BRNs, then we can focus on the biological experiment observations related to interactions between these sub-BRNs. Moreover, this approach corresponds to the classical method used by most biologists when they study a biological system. In fact, they start by studying small BRNs that they believe to be of particular importance in representing a specific biological function. The interactions of this BRN with the external genes, are studied only afterwards even if these external genes potentially could influence the behavior of the studied part. Of course, this bottom-up approach makes sense only if there is a preservation of sub-system behaviors as this hypothesis has been done in this paper under some conditions. Otherwise, systems can only be studied globally because of the presence of emergent properties, that are properties resting on sub-BRNs that are questioned in the global one (see [1, 2] for an abstract mathematical denotation of emergent properties in complex systems).

Outline

In the paper, we address the issue of behavior preservation through path preservation in dynamic models, seen as transition systems: states are Cartesian products of abstract discrete values of concentration levels of gene products ([5]). We now succinctly describe our contributions in a simplified and even approximative way. Clearly, (paths of) dynamic models defined at the level of a global network G cannot be compared directly to (paths of) dynamic models defined at the level of a sub-network S. Thus, to allow model comparisons along an embedding from S to G, on the one hand, following a logical approach, we associate to any dynamic model M of G, called generically a G-model, a S-model R, also called the reduced model of M. On the other hand, by applying quotient technics, the G-model M can also be viewed as a S-model, denoted Q. Roughly speaking, Q is the quotient model of M obtained by abstracting it accordingly to genes and thresholds relevant for the sub-network S. From these preliminary constructions of two S-models, respectively R and Q, from a G-model M, the issue of behavior preservation consists in studying the presence of similar transitions in Q and R, and to put the results in the context of reduction and quotient operations. We identify two remarkable classes of embedding: strict embeddings are such that no gene outside S regulates genes of S while monotonous embeddings are such that any gene outside S regulating at least two genes of S, either activates or inhibits

all its targets. In the case of strict embeddings, the preservation of transitions and therefore of paths is complete in accordance with the intuition, since genes in S do not receive any new regulation via the embedding. In the case of monotonous embeddings, preservation of transitions must be qualified: for any transition of the reduced model R, there exists a path in M, which gives rise to the same transition in the quotient model Q. From these detailed studies of transition preservation in reduced and quotient models, then we get results of formula preservation : all temporal formulas of CTL* ([14]) with comparisons between gene concentrations as atoms and without the **N**ext operator are preserved for strict embeddings while for monotonous embedding, preserved formulas are those with negation only attached to atoms, with **U**(ntil) as unique temporal operator, and with **E**(xistential) as unique quantifier over paths.

Structure of the paper

In this paper, we first present the multi-valued discrete approach developed by R. Thomas [30] for biological regulatory networks as a logic. We then follow the standard approach for presenting a logic, that is defining a syntax (signatures and formulas) and semantics (models and the satisfaction relation). In order to ease the read of the paper, we propose to only present in Section 2 signatures, signature embedding and models that is enough to address our path preservation results along embedding in Section 3. Then, we define formulas over BRN signatures and their satisfaction in models in Section 4 from which we can establish our formula preservation results in Section 5. In Section 6, we will discuss the results in the context of related works. Finally in Section 7 we give some concluding remarks.

2. BRN: syntax, embedding and semantics

2.1. Signatures

A biological regulatory network is represented by a labeled directed graph, called interaction graph. Vertices abstract biological entities, as genes or proteins, and are called variables. Edges abstract interactions between variables. When a variable i can activate a variable j, then there exists an edge from i to j labeled by the sign "+". On the contrary, when a variable i can inhibit a variable j, then there exists an edge from i to j labeled by the sign "-". Moreover, the action, activation or inhibition, between two variables becomes effective only when the concentration of the regulator reaches a given threshold. Thus, each interaction $i \longrightarrow j$ is labeled by a sign and a threshold. In R. Thomas's discrete modeling framework, the concentration levels for the variable i can take a finite number of values $\{0, 1, \ldots, b_i\}$ and thresholds related to the actions of i are labeled with an integer between 1 and b_i. The knowledge of interactions between variables, including signs and these qualitative thresholds, is called the static part of BRNs and constitutes the elements of signatures for a logic dedicated to BRNs.

Definition 1 (Signature). *A BRN-signature G is a labeled directed graph $\langle V, F, Sn, Th \rangle$ where:*

1. *V is a finite set whose elements are called variables.*
2. *$F \subseteq V \times V$ denotes the set of edges.*
3. *Sn is a mapping from F to $\{+, -\}$.*
4. *Th is a mapping from F to $\mathbb{N}^*$ such that:*

$$\forall (i,j) \in F, \ Th(i,j) > 1 \Rightarrow (\ \exists (i,k) \in F, Th(i,k) = Th(i,j) - 1)$$

Point 4. of Definition 1 gives some restrictions on the way the edges are labeled. If an edge outgoing from a variable i is labeled by an integer $l \geq 2$, then there exist edges outgoing from i labeled by $1, \ldots, l-1$. Thus, all intermediate values are used as thresholds and this well represents their qualitative nature.

Remark 1. *Let us emphasize on the particularity of the logic for BRN presented in this paper : signatures are not simple sets of symbols but are interaction graphs (the static part of BRN). This is what makes tough the definition of the embedding (see Definitions 2 and 3) as well as the definitions of the consequences of the embedding for both biological experiment observations expressed over sub-BRNs (see Definition 13) and the dynamics of sub-BRNs embedded into a larger one (see Definition 8). This will also lead to not obvious proofs of our preservation results through embeddings.*

Notation 1.   1. *Let G be a BRN-signature as above and let i be a variable in V. G_i^+ (resp. G_i^-) denotes the set of successors (resp. predecessors) of i in G, and b_i denotes the cardinal of the set of thresholds for i. Formally, we have:*
   - *$G_i^+ = \{ j \in V | (i,j) \in F \}$*
   - *$G_i^- = \{ j \in V | (j,i) \in F \}$*
   - *$b_i = |\{ Th(i,j) \mid j \in G_i^+ \}|$*

2. *In the following, when we consider two BRN-signatures G_1 and G_2, the vertices, edges, signs and thresholds of G_1 and G_2 will be denoted respectively by V_1, F_1, Sn_1, Th_1 and V_2, F_2, Sn_2, Th_2 to make clear the underlying BRN-signature.*

In the rest of this paper, the running examples used to illustrate our approach are purely toy examples and thus, do not represent BRN issued from real case studies. In particular, they are sufficiently small to allow us to draw their models.

Example 1. *Figure 1 shows a network consisting of two genes i and j. More specifically, i activates its own expression above threshold 1, and inhibits the expression of gene j above the same threshold 1, whereas j inhibits the expression of gene i above threshold 1, and the expression of its own gene above threshold 2. The associated BRN-signature, denoted G_1 in the sequel, is given by:*

$$\langle \quad V_1 = \quad \{i, j\},$$
$$F_1 = \quad \{(i,i), (i,j), (j,i), (j,j)\},$$
$$Sn_1 = \quad \{(i,i) \mapsto +, (i,j) \mapsto -, (j,i) \mapsto -, (j,j) \mapsto -\},$$
$$Th_1 = \quad \{(i,i) \mapsto 1, (i,j) \mapsto 1, (j,i) \mapsto 1, (j,j) \mapsto 2\} \quad \rangle$$
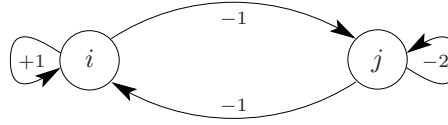
FIGURE 1. BRN-signature G_1 for the $i - j$ network.

2.2. Signature embedding

Biologists can identify small parts issued from a BRN involving a large number of genes. These parts can be considered as a biological function insofar as it can be proven that the biological function is essentially related to the concentration levels of the variables occurring in the considered subpart. Thus, an interesting part of BRN is studied even if it is embedded into the global BRN which controls all cellular processes.

Embedding of BRN-signatures can formalize such an approach. However, a signature embedding cannot be a simple graph embedding. Indeed, as the thresholds on edges depend on the properties of the graph (the threshold of the edge $i \to j$ cannot be greater than the number of edges outgoing from i), it matters to pay attention to the preservation of the conditions on the thresholds (Conditions 3 and 4 of Definition 2). In fact, as thresholds are taken into consideration in signatures, the key point to take into account through the embedding is the preservation of the equality between thresholds and the numerical order between them. New intermediate thresholds for a given variable can be introduced when embedding a BRN into another one, but relationships between existing thresholds have to be preserved in the larger one.

Definition 2 (Signature embedding). *Let* $G = \langle V, F, Sn, Th \rangle$ *and* $G' = \langle V', F', Sn', Th' \rangle$ *be two BRN-signatures. We say that* G *is embedded into* G', *and we note* $G \hookrightarrow G'$, *if:*

1. $V \subseteq V'$ *and* $F = F' \cap V \times V$. *(graph embedding)*
2. $\forall (i, j) \in F,\ Sn(i, j) = Sn'(i, j)$. *(preservation of signs)*
3. $\forall i \in V, \forall j, k \in G_i^+,\ Th(i, j) = Th(i, k) \Leftrightarrow Th'(i, j) = Th'(i, k)$. *(preservation of equality of thresholds)*
4. $\forall i \in V, \forall j, k \in G_i^+,\ Th(i, j) < Th(i, k) \Leftrightarrow Th'(i, j) < Th'(i, k)$. *(preservation of order between thresholds)*

Example 2. *Figure 2 presents a BRN-signature* G_2 *that shares with* G_1 *the two variables* i *and* j, *and introduces the new variable* k. *It is easy to see that* $G_1 \hookrightarrow G_2$. *In the sequel, this embedding will be simply denoted* $\sigma_{12} : G_1 \hookrightarrow G_2$: *the index 12 is a mnemonic sign that indicates that the source and target networks are indexed respectively by 1 and by 2.*

Conditions 1 and 2 are clearly verified since all edges of G_2 *built over vertices of* G_1 *are also in* G_1 *labeled with the same signs.*
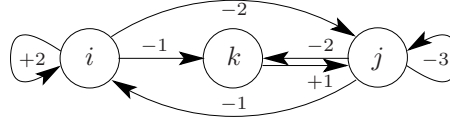
FIGURE 2. BRN-signature G_2.

Condition 3 requires that the equality between thresholds for outgoing edges in G_1 is preserved in G_2. In G_1, we only have $Th_1(i,i) = Th_1(i,j)$ which is preserved in G_2 since $Th_2(i,i) = Th_2(i,j)$.

Finally, condition 4 requires that the order between thresholds for outgoing edges in G_1 is preserved in G_2. For instance, in G_1, j has two outgoing edges (j,i) and (j,j) with $Th_1(j,i) < Th_1(j,j)$. In G_2, we have also $Th_2(j,i) < Th_2(j,j)$.

We will show in Section 5 that the way embeddings are defined in Definition 2 does not allow us to state general results of property preservation. To preserve BRN properties along embeddings (see Section 5), we have to impose supplementary constraints on embeddings.

Definition 3 (Strict and monotonous signature embeddings). *The embedding $G \hookrightarrow G'$ is said*

- *strict if*
$$\forall (i,j) \in F', \qquad j \in V \Rightarrow i \in V.$$

- *monotonous[1] if*
$$\forall i \in V' \setminus V, \ \forall j,k \in V, \quad (i,j) \in F' \wedge (i,k) \in F' \Rightarrow Sn'(i,j) = Sn'(i,k).$$

Remark 2. *Any strict embedding is monotonous.*

Thus, a strict embedding means there is no new regulation on variables of embedded BRN, whereas monotonous embeddings may contain new regulations on embedded variables provided that an external variable has always the same effect (inhibition or activation) on these variables.

Example 3. *The embedding $\sigma_{12} : G_1 \hookrightarrow G_2$ of Example 2 is not strict because of the entering edge from k toward the variable j. Nevertheless, because k has only one outgoing edge, this embedding is monotonous.*

For the BRN-signature G_3 given in Figure 3 which is obtained from G_2 by removing this edge, the signature embedding $\sigma_{13} : G_1 \hookrightarrow G_3$, is strict.

To sum up, in the sequel, we use these two embeddings as running examples to illustrate relationships between the nature of the embedding and the preservation results:

- *$\sigma_{12} : G_1 \hookrightarrow G_2$ is a monotonous embedding (but not a strict embedding)*
- *$\sigma_{13} : G_1 \hookrightarrow G_3$ is a strict embedding (and thus also a monotonous embedding)*

---

[1]$V' \setminus V$ is the set of elements of V' which do not belong to V: $\{x \mid x \in V' \text{ and } x \notin V\}$.
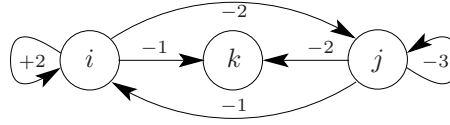
FIGURE 3. BRN-signature G_3.

2.3. Semantics

Models. Each variable i in a BRN-signature G is a genetic entity which is characterized at a given point in time by a concentration level. Since the set of variables and the set of thresholds are finite, the state space generated from G is finite and defined by:

Definition 4 (States). *Let G be a BRN-signature. The state space S_G of G is the set of mappings $s : V \to \mathbb{N}$ such that for every $i \in V$, $s(i) \in \{0, \ldots, b_i\}$.*

Example 4. *In the BRN-signature G_1 of Example 1, variables i and j have respectively 2 and 3 possible concentration levels: 0 or 1, and 0, 1 or 2.*

Therefore, the state space for G_1 is $S_{G_1} = \{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2)\}$ where a state is encoded as an ordered couple of the form $(s(i), s(j))$.

The concentration level of each variable of a given BRN-signature G evolves over time depending on the concentration level of its resources i.e. the sets of i's predecessors in G, corresponding either to activators above their corresponding thresholds, or to inhibitors below their corresponding thresholds. Hence, a resource is the presence of an activator or the absence of an inhibitor. However, neither G nor the concentration level of i's resources give clues to decide how the level of i evolves. This is a degree of freedom of BRN-signatures which gives rise to a class of possible G-models, so-called dynamics of G. However, most of the models defining the dynamics of G, do not describe biological realities. To choose the ones describing such biological realities, some logical approaches based on model-checking technics have been used (e.g. see [5]). Model-checking allows one to choose interesting models satisfying a set of biological experiments described by temporal properties.

Formally, resources are defined as follows:

Definition 5 (Resources). *Let G be a BRN-signature. The set of resources $R_{G,i}(s)$ of a variable i at state $s \in S_G$ is defined by:*

$$R_{G,i}(s) = \begin{cases} \{j \in G_i^- | (Sn(j, i) = + \text{ and } s(j) \geq Th(j, i))\} \\ \cup \\ \{j \in G_i^- | (Sn(j, i) = - \text{ and } s(j) < Th(j, i))\} \end{cases}$$

Example 5. *Figure 4 gives the sets of resources for the variables (i, j and k) for each state of S_{G_1}, S_{G_2} and S_{G_3} respectively.*

Table of resources for G_1

i	j	$R_{G_1,i}$	$R_{G_1,j}$
0	0	$\{j\}$	$\{i,j\}$
0	1	$\emptyset$	$\{i,j\}$
0	2	$\emptyset$	$\{i\}$
1	0	$\{i,j\}$	$\{j\}$
1	1	$\{i\}$	$\{j\}$
1	2	$\{i\}$	$\emptyset$

Table of resources for G_3

i	j	k	$R_{G_3,i}$	$R_{G_3,j}$	$R_{G_3,k}$
0	0	0	$\{j\}$	$\{i,j\}$	$\{i,j\}$
0	1	0	$\emptyset$	$\{i,j\}$	$\{i,j\}$
0	2	0	$\emptyset$	$\{i,j\}$	$\{i\}$
0	3	0	$\emptyset$	$\{i\}$	$\{i\}$
1	0	0	$\{j\}$	$\{i,j\}$	$\{j\}$
1	1	0	$\emptyset$	$\{i,j\}$	$\{j\}$
1	2	0	$\emptyset$	$\{i,j\}$	$\emptyset$
1	3	0	$\emptyset$	$\{i\}$	$\emptyset$
2	0	0	$\{i,j\}$	$\{j\}$	$\{j\}$
2	1	0	$\{i\}$	$\{j\}$	$\{j\}$
2	2	0	$\{i\}$	$\{j\}$	$\emptyset$
2	3	0	$\{i\}$	$\emptyset$	$\emptyset$

Table of resources for G_2

i	j	k	$R_{G_2,i}$	$R_{G_2,j}$	$R_{G_2,k}$
0	0	0	$\{j\}$	$\{i,j\}$	$\{i,j\}$
0	0	1	$\{j\}$	$\{i,j,k\}$	$\{i,j\}$
0	1	0	$\emptyset$	$\{i,j\}$	$\{i,j\}$
0	1	1	$\emptyset$	$\{i,j,k\}$	$\{i,j\}$
0	2	0	$\emptyset$	$\{i,j\}$	$\{i\}$
0	2	1	$\emptyset$	$\{i,j,k\}$	$\{i\}$
0	3	0	$\emptyset$	$\{i\}$	$\{i\}$
0	3	1	$\emptyset$	$\{i,k\}$	$\{i\}$
1	0	0	$\{j\}$	$\{i,j\}$	$\{j\}$
1	0	1	$\{j\}$	$\{i,j,k\}$	$\{j\}$
1	1	0	$\emptyset$	$\{i,j\}$	$\{j\}$
1	1	1	$\emptyset$	$\{i,j,k\}$	$\{j\}$
1	2	0	$\emptyset$	$\{i,j\}$	$\emptyset$
1	2	1	$\emptyset$	$\{i,j,k\}$	$\emptyset$
1	3	0	$\emptyset$	$\{i\}$	$\emptyset$
1	3	1	$\emptyset$	$\{i,k\}$	$\emptyset$
2	0	0	$\{i,j\}$	$\{j\}$	$\{j\}$
2	0	1	$\{i,j\}$	$\{j,k\}$	$\{j\}$
2	1	0	$\{i\}$	$\{j\}$	$\{j\}$
2	1	1	$\{i\}$	$\{j,k\}$	$\{j\}$
2	2	0	$\{i\}$	$\{j\}$	$\emptyset$
2	2	1	$\{i\}$	$\{j,k\}$	$\emptyset$
2	3	0	$\{i\}$	$\emptyset$	$\emptyset$
2	3	1	$\{i\}$	$\{k\}$	$\emptyset$

FIGURE 4. Resources of i, j and k in G_1, G_2 and G_3.

Definition 6 (G-models). *Let G be a BRN-signature.*
Let $\kappa = \{(i,W) \mid i \in V \wedge W \subseteq G_i^-\}$ be the set of all couples (i,W) where i is a variable of G and W a subset of predecessors of i in G.

A G-model is a mapping $p : \kappa \to \mathbb{N}$ such that for every (i,W) in κ, we have $p(i,W) \in \{0, \ldots, b_i\}$.

Intuitively, the mapping p applied to (i,W) gives the concentration level towards which the variable i tends to evolve when W corresponds to its set of resources.

Example 6. *From the BRN-signature G_2 of Figure 2, we have the following set κ:*

$$\kappa = \begin{cases} \{(i,\emptyset), (i,\{i\}), (i,\{j\}), (i,\{i,j\})\} \\ \cup \\ \{(j,\emptyset), (j,\{i\}), (j,\{j\}), (j,\{k\}), (j,\{i,j\}), (j,\{i,k\}), (j,\{j,k\}), (j,\{i,j,k\})\} \\ \cup \\ \{(k,\emptyset), (k,\{i\}), (k,\{j\}), (k,\{i,j\})\} \end{cases}$$

A possible G_2-model p_2 is given in Figure 5 (left).

Signature embeddings $\sigma : G \hookrightarrow G'$ (see Definition 2) have a counterpart on models which is expressed by a classic forgetful mapping from the "richer" model defined on G' to the "poorer" one defined on G [20]. Some difficulties occur due to some restrictions to make on thresholds. These restrictions are defined by a V-indexed family of mapping σ_i that are largely used in the rest of the paper. They are defined as follows:

Definition 7 (Threshold mappings). *Let $G = \langle V, F, Sn, Th \rangle$ and $G' = \langle V', F', Sn', Th' \rangle$ be two BRN-signatures such that G is embedded into G' by means of the signature embedding: $\sigma : G \hookrightarrow G'$.*

For every $i \in V$, let us define the mapping $\sigma_i : \{0, 1, \ldots, b_i\} \to \{0, 1, \ldots, b'_i\}$ as follows:

- *$\sigma_i(0) = 0$.*
- *For every $l \neq 0$, $\sigma_i(l) = Th'(i, j)$ with j any variable in G_i^+ such that $Th(i, j) = l$.*

We have the following obvious result:

Fact 1. *σ_i is strictly increasing: if $l < l'$ then $\sigma_i(l) < \sigma_i(l')$.*

Proof. This is obvious if $l = 0$. Otherwise, by the definition of a signature, there exists (i, j) and (i, k) in F such that $l = Th(i, j)$ and $l' = Th(i, k)$. Then, $\sigma_i(l) = Th'(i, j)$ and $\sigma_i(l') = Th'(i, k)$ and it is then sufficient to remark that, by the definition of a signature embedding, if $Th(i, j) < Th(j, k)$ then $Th'(i, j) < Th'(i, k)$. □

It is now possible to define the reduced model of any model of the "richer" signature.

Definition 8 (Reduced model). *Let G and G' be two BRN-signatures such that there exists a signature embedding $\sigma : G \hookrightarrow G'$ and let p' be a G'-model. The reduced G-model p from p' is the G-model defined as follows:*

$$\forall i \in V, \forall W \subseteq G_i^-, \quad p(i, W) = \max\{l \in \{0, \ldots, b_i\} \mid \sigma_i(l) \leq p'(i, W)\}$$

Example 7. *Figure 5 and Figure 6 give, respectively, the reduced G_1-model from p_2, denoted p_{12}, and the reduced G_1-model from p_3, denoted p_{13}, along respectively the signature embeddings $\sigma_{12} : G_1 \hookrightarrow G_2$ and $\sigma_{13} : G_1 \hookrightarrow G_3$ introduced in Example 3.*

resources W_2	$p_2(i, W_2)$	$p_2(j, W_2)$	$p_2(k, W_2)$
∅	0	0	0
{i}	2	2	1
{j}	2	1	1
{k}		1	
{i,j}	2	2	1
{i,k}		3	
{j,k}		3	
{i,j,k}		3	

resources W_1	$p_{12}(i, W_1)$	$p_{12}(j, W_1)$
∅	0	0
{i}	1	1
{j}	1	1
{i,j}	1	1

FIGURE 5. A G_2-model p_2 (left) and its reduced G_1-model p_{12} (right).

resources W_3	$p_3(i, W_3)$	$p_3(j, W_3)$	$p_3(k, W_3)$
∅	0	0	0
{i}	2	2	0
{j}	2	1	0
{i,j}	2	3	0

resources W_1	$p_{13}(i, W_1)$	$p_{13}(j, W_1)$
∅	0	0
{i}	1	1
{j}	1	1
{i,j}	1	2

FIGURE 6. A G_3-model p_3 (left) and its reduced G_1-model p_{13} (right).
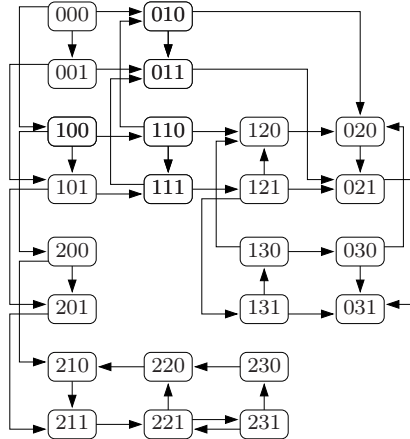
FIGURE 7. Asynchronous transition system $ATS(G_2, p_2)$.

From a G-model p, a transition system (S_G, T) can be generated where the transitions in T give the state evolution as described in p. Here, there are two possibilities: either the concentration level of two or more variables can change simultaneously towards their corresponding target concentration level specified by p, or the change of a value is always done stepwise (increasing or decreasing by 1), that is only one of the involved variables can be modified by one unit, at a time. These two possibilities are respectively called synchronous and asynchronous description of the G-model p. Here, we take the asynchronous description because it is unlikely that, in vivo, several variables cross their thresholds simultaneously or a single variable crosses several thresholds simultaneously [29].

Definition 9 (Asynchronous transition system). *Let G be a BRN-signature and let p be a G-model. The Asynchronous Transition System generated from p is the directed graph $ATS(G, p) = (S_G, T)$ defined by:*

  $\forall s, s' \in S_G,\ (s, s') \in T$ *if, and only if:*
    − *there exists $i \in V$, such that*

$$s'(i) = \begin{cases} s(i) \ + \ 1 \ and \ s(i) \ < p(i, R_{G,i}(s)) \\ s(i) \ - \ 1 \ and \ s(i) \ > p(i, R_{G,i}(s)) \end{cases}$$

    − *and $s'(j) = s(j)$ for every $j \in V \setminus \{i\}$.*

Example 8. *Figure 7 gives the asynchronous transition system $ATS(G_2, p_2)$ associated to the G_2-model p_2.*

3. Path preservation results

In this section, we consider two BRN-signatures G and G' such that $\sigma : G \hookrightarrow G'$. Hence, given a G'-model p' and the reduced G-model p obtained from p', we

study the relationships between the asynchronous transition systems $ATS(G',p')$ and $ATS(G,p)$. We define an equivalence relation $\simeq$ on the states of G' leading to a quotient transition system and then, show how the paths of $ATS(G,p)$ are preserved in $ATS(G',p')$ up to this quotient. Before establishing these preservation results, let us introduce first some notions and notations about path and quotient of models over a BRN signature.

3.1. Paths and quotients of transition systems

Let G be a BRN signature and let p be a G-model.

Path. A path in $ATS(G,p) = (S_G, T)$ is a non empty finite or infinite sequence $(s_0, s_1, \ldots)$ such that $(s_i, s_{i+1}) \in T$ for all $i \geq 0$. For a state $s \in S_G$, the set of paths $(s_0, s_1, \ldots)$ starting from s, that is verifying $s_0 = s$ is denoted as $path(s)$. For a path $\rho = (s_0, s_1, \ldots)$, s_0 is noted $first(\rho)$.

For a finite or infinite path $\rho = (s_0, s_1, \ldots)$, for an index i such that s_i belongs to ρ, we note ρ^i the non empty path $(s_i, \ldots)$ starting from s_i and containing all states s_j with $i \leq j$.

$Index(\rho)$ is the set of indexes i such that ρ^i is defined. We then have the equality $s_i = first(\rho^i)$ for any i in $Index(\rho)$.

Any path of the form ρ^i is called a suffix of ρ. We note $\rho \leq \theta$ (resp. $\rho < \theta$) when θ is a suffix of ρ (resp. θ is a proper suffix of ρ, i.e. a path ρ^i with $i > 0$).

Quotient. Let $\simeq$ be an equivalence relation on S_G. The quotient of (S_G, T) by $\simeq$, denoted $(S_G, T)_{/\simeq} = (S_{/\simeq}, T_{/\simeq})$, is defined as follows:

- the set of states $S_{G/\simeq}$ is the set of equivalence classes of $\simeq$, $[s]$ denoting the equivalence class of a state s of S_G;
- the set of transitions $T_{/\simeq}$ is defined by $([s],[t]) \in T_{/\simeq}$ iff there exists $s' \in [s]$ and $t' \in [t]$ such that $(s', t') \in T$.

3.2. Partition

First, we introduce the partition defined on the set of states of G' that allows us to define the quotient transition system. In the next definition, we introduce a mapping B from S_G to the set of subsets of $S_{G'}$, that we will use to define the equivalence relation $\simeq$.

Definition 10 (Partition). *We define the mapping $B : S_G \to 2^{S_{G'}}$ as follows. For every $s \in S_G$, $B(s)$ is the set of states s' in $S_{G'}$ such that for every i in V and using the family of threshold mappings $\sigma_i : \{0, 1, \ldots, b_i\} \to \{0, 1, \ldots, b'_i\}$ introduced in Definition 7:*

$$\sigma_i(s(i)) \leq s'(i) < \sigma_i(s(i) + 1)$$

if $s(i) < b_i$, and

$$\sigma_i(s(i)) \leq s'(i)$$

otherwise.

The interest of this mapping rests on the two following results:

Proposition 1. *The mapping B makes a partition of $S_{G'}$.*

Proof. We first prove that, for all $s_1, s_2 \in S_G$, if $s_1 \neq s_2$ then $B(s_1) \cap B(s_2) = \emptyset$. Indeed, if $s_1 \neq s_2$ there exists $i \in V$ such that $s_1(i) \neq s_2(i)$. Without loss of generality, suppose that $s_1(i) < s_2(i)$. Then, $s_1(i) + 1 \leq s_2(i)$ and by Fact 1,

$$\sigma_i(s_1(i) + 1) \leq \sigma_i(s_2(i))$$

If $s' \in B(s_1)$ then

$$s'(i) < \sigma_i(s_1(i) + 1) \leq \sigma_i(s_2(i))$$

so $s' \notin B(s_2)$, and if $s' \in B(s_2)$ then

$$\sigma_i(s_1(i) + 1) \leq \sigma_i(s_2(i)) \leq s'(i)$$

so $s' \notin B(s_1)$. Thus $B(s_1) \cap B(s_2)$ is actually empty. It remains to prove that

$$\cup_{s \in S_G} B(s) = S_{G'}$$

Since $B(s) \subseteq S_{G'}$ for all $s \in S_G$, it is in fact sufficient to observe that, given any state s' in $S_{G'}$, we have $s' \in B(s)$ for the state $s \in S_G$ defined by, for every $i \in V$,

$$s(i) = \max\{l \in \{0, \ldots, b_i\} \mid \sigma_i(l) \leq s'(i)\}$$

$\qquad\square$

Proposition 2. *For every $s \in S_G$, every $s' \in B(s)$ and every $(r, i) \in F$:*

$$r \in R_{G,i}(s) \iff r \in R_{G',i}(s')$$

Proof. Since $Sn(r, i) = Sn'(r, i)$, it is sufficient to prove that

$$Th(r, i) \leq s(r) \iff Th'(r, i) \leq s'(r)$$

Suppose that

$$Th(r, i) \leq s(r)$$

Then, from the definition of σ_r, Fact 1 and $s' \in B(s)$, we deduce that

$$Th'(r, i) = \sigma_r(Th(r, i)) \leq \sigma_r(s(r)) \leq s'(r)$$

Now, suppose that $Th(r, i) > s(r)$. Then

$$Th(r, i) \geq s(r) + 1$$

and from the definition of σ_r, Fact 1 and $s' \in B(s)$, we deduce that

$$Th'(r, i) = \sigma_r(Th(r, i)) \geq \sigma_r(s(r) + 1) > s'(r)$$

$\qquad\square$

Following Proposition 1, we can define the equivalence relation $\simeq$ on $S_{G'}$ by:

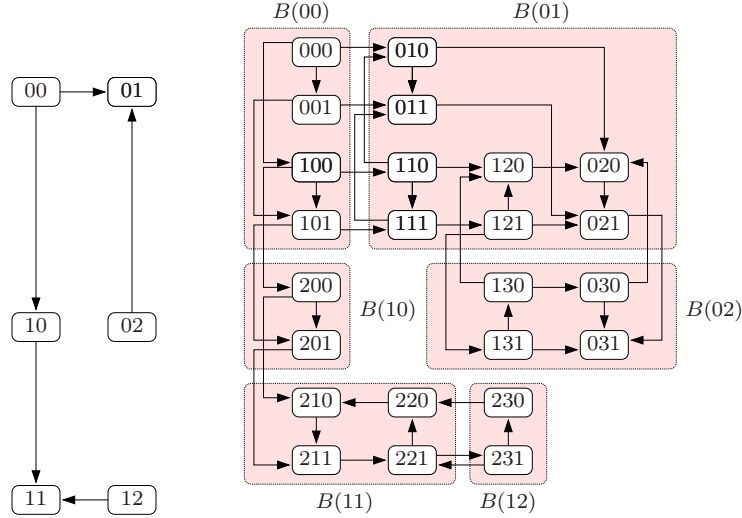$$s_1' \simeq s_2' \iff B^{-1}(s_1') = B^{-1}(s_2')$$

FIGURE 8. Asynchronous transition systems $ATS(G_1, p_{12})$ (left) and $ATS(G_2, p_2)$ (right).
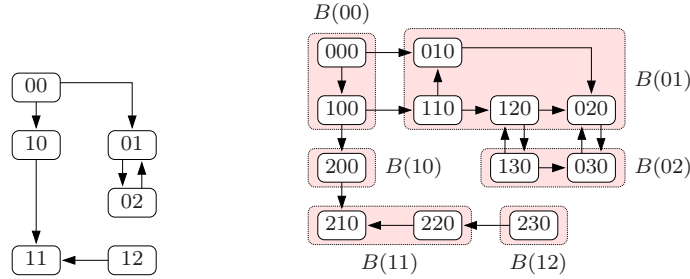

FIGURE 9. Asynchronous transition systems $ATS(G_1, p_{13})$ (left) and $ATS(G_3, p_3)$ (right).

Example 9. *On the right part of respectively Figures 8 and 9, we give the asynchronous transition systems $ATS(G_2, p_2) = (S_{G_2}, T_2)$ and $ATS(G_3, p_3) = (S_{G_3}, T_3)$ associated respectively to p_2 and p_3. ($ATS(G_2, p_2)$ has already been introduced in Example 7.) States are represented by white boxes that are labelled by the tuple of concentration level values of the involved variables: for example, 010 or 220 are such labels.*

In both figures, colored boxes contain several white boxes (states) and represent the corresponding $\simeq$-equivalence classes. They are labelled by names of the form $B(s(i)s(j))$ making clear the relationship between the equivalence class and the corresponding state of S_{G_1} where G_1 is the source signature of the considered signature embeddings, respectively $\sigma_{12} : G_1 \hookrightarrow G_2$ and $\sigma_{13} : G_1 \hookrightarrow G_3$.

Given a G'-model p' and its reduced G-model p, we can now study the relationships between asynchronous transition systems $ATS(G', p')$ and $ATS(G, p)$. In the following, we show two results depending on whether the embedding is monotonous or strict. Hence,

- we prove that if the considered embedding $G \hookrightarrow G'$ is monotonous, and if $(s_0, s_1, \ldots, s_n)$ is a path of $ATS(G, p)$, then $ATS(G', p')$ contains a path going successively through the equivalence classes $B(s_0), B(s_1), \ldots, B(s_n)$. Observe, for example in Figure 8 that, for the path $(02), (01)$ in $ATS(G_1, p_{12})$, we associate the path $(130), (030), (020), (021)$ in $ATS(G_2, p_2)$ going through the equivalence classes $B(02), B(01)$. This implies that each finite path of $ATS(G, p)$ is preserved in the quotient transition system $ATS(G', p')_{/\simeq}$.
- Then, we show a stronger result when the embedding is strict: first, we will show that, if $(s_0, s_1, \ldots, s_n)$ is a path of $ATS(G, p)$ then $ATS(G', p')$ contains, for all $s' \in B(s_0)$, a path starting from s' which goes successively through the classes $B(s_0), B(s_1), \ldots, B(s_n)$. Figure 9, for example, shows that for the path $(00), (01), (02)$ in $ATS(G_1, p_{13})$, and for any state in $B(00)$, we can find a path in $ATS(G_3, p_3)$ going through the equivalence classes $B(00), B(01), B(02)$. We will show that transitions between different states in $ATS(G, p)$ are preserved in $ATS(G', p')_{/\simeq}$ and vice versa. Hence, there is a complete preservation of dynamics of any sub-system when dealing with strict embeddings.

3.3. Path preservation for monotonous embedding

If the embedding $G \hookrightarrow G'$ is monotonous, then any variable $i \in V'$ that does not belong to G acts on the variables of G according to a same sign. If this sign is positive (resp. negative), then i is a resource of no variable of G whenever its level is minimal (resp. maximal). In other words, it is always possible to fix the level of the variables acting on G so that all the corresponding interactions on G "vanish". The sub-BRN signature G is then "isolated" from the larger signature G' in which it is embedded. This key idea will be further used to prove the preservation results in the rest of this section. Intuitively, these results show that the behaviors of the variables of G described by the G'-model p' are, under the "isolation" mentioned above, strongly related to the behaviors of the variables of G described by the reduced G-model obtained from p'.

Let us first define, for each $s \in S_G$, a subset $B'(s)$ of $B(s)$ in which the variables acting on G are resources of no variable of G (so that in each state of $B'(s)$, the "isolation" mentioned above occurs).

Definition 11. *Suppose that the embedding $G \hookrightarrow G'$ is monotonous. We define the mapping $B' : S_G \to 2^{S_{G'}}$ as follows: for every $s \in S_G$, $B'(s)$ is the set of $s' \in B(s)$ such that for $i \in V' \setminus V$ regulating at least a variable of V,*

$$s'(i) = \begin{cases} 0 & \text{if } \exists j \in V, Sn'(i, j) = + \\ b'_i & \text{if } \exists j \in V, Sn'(i, j) = - \end{cases}$$

Let us point out that such $B'(s)$ sets are defined only for monotonous embeddings. Indeed, if there exists $i \in V' \setminus V$, with at least two variables j_1 and j_2 in V such that $Sn'(i, j_1) = +$ and $Sn'(i, j_2) = -$, then no value can be attached to $s'(i)$ in $B'(s)$. Hence, each set $B'(s)$ is not empty by construction.

Remark 3. *If the embedding $G \hookrightarrow G'$ is strict, then the embedding is also monotonous. $B'(s)$ is defined for all $s \in S_G$ and verifies the equality $B'(s) = B(s)$.*

Lemma 1. *For every $s \in S_G$, every $s' \in B'(s)$ and every $i \in V$, we have $R_{G',i}(s') = R_{G,i}(s)$.*

Proof. According to Proposition 2, we have

$$R_{G,i}(s) = R_{G',i}(s') \cap V \tag{1}$$

It is thus sufficient to prove that $R_{G',i}(s') \subseteq V$. So let $r \in R_{G',i}(s')$. First, if $Sn'(r, i) = +$ then $s'(r) \geq Th'(r, i) > 0$, and since $s' \in B'(s)$, we deduce that $r \in V$. Then, if $Sn'(r, i) = -$, we have $s'(r) < Th'(r', i) \leq b'_r$, and since $s' \in B'(s)$, we deduce that $r \in V$. $\square$

Lemma 2. *Let p' be a G'-model and let p be the reduced G-model from p'. For every $s \in S_G$, every $s' \in B'(s)$ and every $i \in V$,*

$$s(i) < p(i, R_{G,i}(s)) \Rightarrow s'(i) < p'(i, R_{G',i}(s'))$$

and

$$s(i) > p(i, R_{G,i}(s)) \Rightarrow s'(i) > p'(i, R_{G',i}(s'))$$

Proof. Suppose that $s(i) < p(i, R_{G,i}(s))$, that is,

$$s(i) + 1 \leq p(i, R_{G,i}(s))$$

By Fact 1 and the definition of p,

$$\sigma_i(s(i) + 1) \leq \sigma_i(p(i, R_{G,i}(s))) \leq p'(i, R_{G,i}(s))$$

We deduce from the Lemma 1 that

$$\sigma_i(s(i) + 1) \leq \sigma_i(p(i, R_{G,i}(s))) \leq p'(i, R_{G',i}(s'))$$

and from $s' \in B(s)$ we obtain

$$s'(i) < \sigma_i(s(i) + 1) \leq \sigma_i(p(i, R_{G,i}(s))) \leq p'(i, R_{G',i}(s'))$$

Now, suppose that

$$p(i, R_{G,i}(s)) < s(i)$$

From the definition of p and $s \in B(s)$ we have

$$p'(i, R_{G,i}(s)) < \sigma_i(s(i)) \leq s'(i)$$

Using Lemma 1 we obtain

$$p'(i, R_{G',i}(s')) < \sigma_i(s(i)) \leq s'(i)$$

$\square$

The main path preservation result for monotonous embedding follows.

Lemma 3. *Suppose that the embedding $G \hookrightarrow G'$ is monotonous. Let p' be a G'-model and let p be the reduced G-model from p'.*

If (s, t) is a transition of $ATS(G, p)$, then for all $s' \in B'(s)$ there exists $t' \in B'(t)$ such that $ATS(G', p')$ has a path from s' to t' whose vertices, except t', belong to $B'(s)$.

Proof. Suppose (s, t) to be a transition of $ATS(G, p)$. By definition, there exists $i \in V$ such that $s(j) = t(j)$ for all $j \in V \setminus \{i\}$ and

$$t(i) = \begin{cases} s(i) + 1 \text{ and } s(i) < p(i, R_{G,i}(s)) \\ \text{or} \\ s(i) - 1 \text{ and } s(i) > p(i, R_{G,i}(s)) \end{cases}$$

Suppose that $s(i) < p(i, R_{G,i}(s))$, the proof being similar in the other case. For all $s' \in B'(s)$, let s'_t be the state of $S_{G'}$ defined by:

$$s'_t(i) = \sigma_i(t(i)) \qquad \text{and} \qquad s'_t(j) = s'(j) \text{ for all } j \in V' \setminus \{i\}$$

It is easy to see that $s'_t \in B'(t)$. Indeed, since $s' \in B'(s)$ and since for every $j \in V' \setminus V$ we have $s'_t(j) = s'(j)$, to prove that $s'_t \in B'(t)$ it is sufficient to prove that $s'_t \in B(t)$. Now, since $s' \in B(s)$, and since for every $j \in V \setminus \{i\}$ we have both $s'_t(j) = s'(j)$ and $t(j) = s(j)$, to prove that $s'_t \in B(t)$, it is sufficient to prove that $\sigma_i(t(i)) \le s'_t(i)$ and, if $t(i) < b_i$, that $s'_t(i) < \sigma_i(t(i) + 1)$. Since $\sigma_i(t(i)) = s'_t(i)$ it is sufficient to remark that, if $t(i) < b_i$, then following Fact 1, $\sigma_i(t(i)) < \sigma_i(t(i) + 1)$.

So to prove the lemma, it is sufficient to prove, for all $s' \in B'(s)$, the presence of a path from s' to s'_t whose vertices, except s'_t, belong to $B'(s)$. We proceed by induction on

$$d(s') = s'_t(i) - s'(i)$$

1. *Base case: $d(s') = 1$.* By Lemma 2,

$$s'(i) < p'(i, R'_{G',i}(s'))$$

   From $d(s') = 1$ and the definition of s'_t, we deduce that (s', s'_t) is a transition of $ATS(G', p')$.

2. *Induction step: $d(s') > 1$.* Let s'' be the state of $S_{G'}$ defined by

$$s''(i) = s'(i) + 1 \quad \text{and} \quad s''(j) = s'(j) \text{ for all } j \in V' \setminus \{i\}$$

   Clearly, $s'' \in B'(s)$. Indeed, since $s' \in B'(s)$ and since for every $j \in V' \setminus V$ we have $s''(j) = s'(j)$, to prove that $s'' \in B'(s)$ it is sufficient to prove that $s'' \in B(s)$. Now, since $s' \in B(s)$, and since for every $j \in V \setminus \{i\}$ we have $s'_t(j) = s'(j)$, to prove that $s'' \in B(s)$, it is sufficient to prove that $\sigma_i(s(i)) \le s''(i) < \sigma_i(s(i) + 1)$ (we have $s(i) < t(i) \le b_i$). This is obvious: since $s' \in B(s)$, $\sigma_i(s(i)) \le s'(i) < s'(i) + 1 = s''(i)$ and, since $d(s') > 1$, $s''(i) = s'(i) + 1 < s'_t(i) = \sigma_i(t(i)) = \sigma_i(s(i) + 1)$.

   Moreover, we have also $d(s'') < d(s')$ and $s''_t = s'_t$. So, by induction hypothesis, there is a path from s'' to s'_t whose vertices except s'_t belongs to $B'(s)$. Then, by Lemma 2,

$$s'(i) < p'(i, R'_{G',i}(s'))$$

and it comes from the definition of s'' that (s', s'') is a transition of $ATS(G', p')$. There is thus a path from s' to s'_t whose vertices except s'_t belong to $B'(s)$.

$\square$

By applying several times the previous lemma, we obtain the following result holding on paths instead of transitions.

Theorem 1. *Suppose that the embedding $G \hookrightarrow G'$ is monotonous, let p' be a G'-model and let p be the reduced G-model from p'.*

If ρ is a path of $ATS(G, p)$, then for every $s' \in B'(first(\rho))$, there exists a path ρ' in $ATS(G', p')$, verifying:

- $first(\rho') = s'$,
- *for all indexes i of $Index(\rho)$, there exists an index k_i in $Index(\rho')$ such that the family $(k_i)_{i \in Index(\rho)}$ is strictly increasing, i.e. $\forall i, i' \in Index(\rho), i < i' \Rightarrow k_i < k_{i'}$*
- $\forall j \in Index(\rho')$,

$$(\exists i, k_i \leq j < k_{i+1} \Rightarrow s'_j \in B'(s_i)) \tag{*}$$

or

$$(\forall i \in Index(\rho), k_i \leq j) \Rightarrow s'_j \in B'(s_{max(Index(\rho))}) \tag{**}$$

where the function $max(E)$ simply gives the largest element of the finite set E.

The case (*) corresponds to nominal indexes of infinite or finite paths while the (**) corresponds to the particular treatment of the last index of a finite path.

We generically note $B'(\rho)^{s'}$ such a path ρ' as characterized in Theorem 1, and we note $B'(\rho)$ the set of all paths $B'(\rho)^{s'}$, with $s' \in B'(first(\rho))$. Thus, we have :

$$B'(\rho) = \{B'(\rho)^{s'} \mid s' \in B'(first(\rho))\}$$

Moreover, each ρ' in $B'(\rho)$ can be provided with an application $Index_\rho$ from $Index(\rho)$ to $\mathbb{N}$, associating to each i in $Index(\rho)$, the integer k_i.

3.4. Path preservation for strict embedding

Here, we give a stronger path preservation result when dealing with strict embedding. This comes from the fact that, for strict embedding, the "isolation" of G occurs not only for states $s' \in B'(s)$, but for all states $s' \in B(s)$ (cf. Remark 3).

Lemma 4. *Suppose that the embedding $G \hookrightarrow G'$ is strict, let p' be a G'-model and let p be the reduced G-model from p'. If (s, t) is a transition of $ATS(G, p)$ then, for all $s' \in B(s)$, there exists $t' \in B(t)$ such that $ATS(G', p')$ has a path from s' to t' whose vertices, except t', belong to $B(s)$.*

Proof. Straightforward from Remark 2, Remark 3 and Lemma 3. $\square$

In addition, for strict embedding, transitions between different states in $ATS(G, p)$ and the quotient of $ATS(G', p')$ by $\simeq$ are preserved.

Theorem 2. *Suppose that the embedding $G \hookrightarrow G'$ is strict. Let p' be a G'-model and let p be the reduced G-model from p'. Then, (s, t) is a transition of $ATS(G, p)$ if and only if $(B(s), B(t))$ is a transition of $ATS(G', p')_{/\simeq}$.*

Proof. If (s, t) is a transition of $ATS(G, p)$, then following Lemma 4, there exists $s' \in B(s)$ and $t' \in B(t)$ such that (s', t') is a transition of $ATS(G', p')$, so that $(B(s), B(t))$ is a transition of $ATS(G', p')_{/\simeq}$. It remains to prove that if $(B(s), B(t))$ is a transition of $ATS(G', p')_{/\simeq}$ then (s, t) is a transition of $ATS(G, p)$. So let $(B(s), B(t))$ be a transition of $ATS(G', p')_{/\simeq}$. There exists $s' \in B(s)$ and $t' \in B(t)$ such that (s', t') is a transition of $ATS(G', p')$. By definition, there exists $i \in V'$ such that $s'(j) = t'(j)$ for all $j \in V' \setminus \{i\}$ and

$$t'(i) = \begin{cases} s'(i) + 1 \text{ and } s'(i) < p'(i, R_{G', i}(s')) \\ s'(i) - 1 \text{ and } s'(i) > p'(i, R_{G', i}(s')). \end{cases}$$

Clearly, $i \in V$ since otherwise s' and t' are in the same equivalence class. Now, suppose that

$$s'(i) < p'(i, R_{G', i}(s')),$$

the proof being similar in the other case. Let $j \in V$. If $t(j) < s(j)$ then

$$t(j) + 1 \le s(j)$$

so, by Fact 1,

$$\sigma_j(t(j) + 1) \le \sigma_j(s(j))$$

and since $s' \in B(s)$ and $t' \in B(t')$ we obtain

$$t'(j) < \sigma_j(t(j) + 1) \le \sigma_j(s(j)) \le s'(j),$$

a contradiction. So, $s(j) \le t(j)$ for all $j \in V$. With similar arguments, we show that $s(j) \ge t(j)$ for all $j \in V \setminus \{i\}$. Consequently,

$$s(i) \le t(i) \quad \text{and} \quad s(j) = t(j) \text{ for all } j \in V \setminus \{i\}$$

Since $B(s) \ne B(t)$ we deduce that $s(i) < t(i)$. So $s(i) + 1 \le t(i)$ and following Fact 1,

$$\sigma_j(s(j) + 1) \le \sigma_j(t(j))$$

Since $s' \in B(s)$ and $t' \in B(t')$ we obtain

$$s'(i) < \sigma_i(s(i) + 1) \le \sigma_i(t(i)) \le t'(i) = s'(i) + 1.$$

Hence,

$$\sigma_i(s(i) + 1) = \sigma_i(t(i))$$

Following Fact 1, σ_i is an injection, so

$$s(i) + 1 = t(i)$$

and to prove that (s, t) is a transition of $ATS(G, p)$ it remains to prove that

$$s(i) < p(i, R_{G, i}(s)) \tag{2}$$

Since

$$s'(i) < p'(i, R_{G', i}(s'))$$

we have
$$t'(i) \leq p'(i, R_{G',i}(s'))$$
and since $t' \in B(t)$ we obtain
$$\sigma_i(t(i)) \leq t'(i) \leq p'(i, R_{G',i}(s'))$$
We then deduce from Remark 2, Remark 3 and Lemma 1 that
$$\sigma_i(t(i)) \leq p'(i, R_{G,i}(s))$$
Following the definition of p, it means that
$$t(i) \leq p(i, R_{G,i}(s))$$
and since $s(i) < t(i)$ the inequality (2) is proved.

$\square$

Theorem 3. *Suppose that the embedding $G \hookrightarrow G'$ is strict, and let p' be a G'-model. Let $s, r \in S_G$, let $s'_1, s'_2 \in B(s)$ and let $r'_1 \in B(r)$ be such that (s'_1, r'_1) is a transition of $ATS(G', p')$. Then there exists $r'_2 \in B(r)$ such that $ATS(G', p')$ has a path from s'_2 to r'_2 whose vertices, except r'_2, belong to $B(s)$.*

Proof. Let p be the reduced G-model from p'. Since (s'_1, r'_1) is a transition of $ATS(G', p')$, $(B(s), B(r))$ is a transition of $ATS(G', p')_{/\simeq}$, and we deduce from Theorem 2 that (s, r) is a transition of $ATS(G, p)$. Then, following Lemma 4, there exists $r'_2 \in B(r)$ such that $ATS(G', p')$ has a path from s'_2 to r'_2 whose vertices, except r'_2, belong to $B(s)$. $\square$

4. BRN logic

In this section, we complete the BRN logic by defining the set of well-formed formulas over BRN signatures, and by giving the definition of their satisfaction for BRN models. This will allow us to address the property preservation through embeddings using our path preservation results given in Section 3.

4.1. Formulas

We propose to express properties over BRN by formulas of the logic CTL*. The logic CTL* combines the expressive power of Computational tree logic (CTL) and Linear temporal logic (LTL) [14]. CTL is a branching-time temporal logic where the structure representing all possible executions is tree-like rather than linear while LTL is a linear temporal logic that implicitly quantifies universally over paths starting from a given state and modelizes linearly all possible executions. Here, we consider actually the restriction of CTL* by removing the next operator X, noted respectively CTL*-X [18, 31]. For biological applications, the time mandatory for a biological system to change of qualitative state, i.e the time for a neXt transition, has a large variance. Thus, an experiment along which a state s' is observed after a state s does not imply that s' is a (direct) successor of state s because it is difficult to know if other states have been visited in between. This explains why the modal

operator X is not of big relevance. Hence, formulas for BRN are simply $CTL^\star$-X formulas whose atomic formulas describe comparisons between a concentration level of a variable with some threshold values. $CTL^\star$-X formulas include *state and path formulas* which are true in a given state and along a given path, respectively. They are well-formed formulas whose the syntactical rules are given below where $SFor$ and $PFor$ stand respectively for state formulas and for path formulas.

Definition 12 (BRN Formulas or $CTL^\star$-X formulas on a BRN). *Let $G = \langle V, F, Sn, Th \rangle$ be a BRN-signature. Formulas over G are defined as follows:*

$$
\begin{aligned}
SFor ::= \quad & Atom(G) \mid \\
& \neg SFor \mid SFor \wedge SFor \mid \\
& \mathbf{E} \ PFor \\
PFor ::= \quad & SFor \mid \\
& \neg PFor \mid PFor \wedge PFor \mid \\
& PFor \ \mathbf{U} \ PFor
\end{aligned}
$$

where $Atom(G)$ are all atomic formulas of the form $(i \sim l)$ where $i \in V$, $l \in \{0, \ldots, b_i\}$ and $\sim \in \{=, <, >\}$.

We denote by $Sen(G)$ the set of formulas over G.

In the sequel, $i \geq l$ (resp. $i \leq l$) denotes the formula $i = l \vee i > l$ (resp. $i = l \vee i < l$).

As usual, we introduce some notations commonly used when dealing with temporal logics: $\mathbf{T}$ for $\neg(at \wedge \neg at)$ with at any arbitrary element of $Atom(G)$, $\varphi \vee \varphi'$ for $\neg(\neg\varphi \wedge \neg\varphi')$, $\varphi \Rightarrow \varphi'$ for $\neg\varphi \vee \varphi'$, $\mathbf{A} \ \varphi$ for $\neg\mathbf{E} \ \neg\varphi$, $\mathbf{F} \ \varphi$ for $\mathbf{T} \ \mathbf{U} \ \varphi$, and $\mathbf{G} \ \varphi$ for $\neg\mathbf{F} \ \neg\varphi$.

The intuitive meaning of all these symbols is: $\mathbf{F} \ \varphi$ (resp. $\mathbf{G} \ \varphi$, $\varphi \ \mathbf{U} \ \psi$) for: φ will be finally ($\mathbf{F}$) true (resp. is globally ($\mathbf{G}$) true, φ has to be true until ($\mathbf{U}$) ψ becomes true). The prefix $\mathbf{A}$ (resp. $\mathbf{E}$) means that the formula is true for all possible futures represented by paths issued from the current state (resp. for at least one future).

In the context of a signature embedding $G \hookrightarrow G'$, formulas resting on G have to be adapted because of thresholds occurring in atomic formulas. Indeed, the presence of new variables makes side effects on the thresholds by shifting them. The following definition expresses how to convert formulas in $Sen(G)$ into formulas in $Sen(G')$ by following the simple idea of translating a threshold into an interval of possible values.

Definition 13 (Formula renaming). *Let G and G' be two BRN-signatures such that G is embedded into G'. Let us note $\sigma : Atom(G) \to Sen(G')$ the mapping defined by:*

- *For all $(i = l) \in Atom(G)$ with $l \neq b_i$, $\sigma(i = l)$ is $i \geq \sigma_i(l) \wedge i < \sigma_i(l + 1)$.*
- *For all $(i = b_i) \in Atom(G)$, $\sigma(i = b_i)$ is $i \geq \sigma_i(b_i)$.*
- *For all $(i > l) \in Atom(G)$, $\sigma(i > l)$ is $i \geq \sigma_i(l + 1)$.*

- *For all $(i < l) \in Atom(G)$, $\sigma(i < l)$ is $i < \sigma_i(l)$.*

Let us note $\bar{\sigma}$ the canonical extension of $\sigma : Sen(G) \to Sen(G')$ on formulas in $Sen(G)$ defined as follows:

- *For $at \in Atom(G)$, $\bar{\sigma}(at) = \sigma(at)$,*
- *For other formulas, Boolean connectives and temporal operators are preserved.*

Example 10. *For the embedding $\sigma_{12} : G_1 \hookrightarrow G_2$ of Example 2, we have $\sigma_{12i}(1) = 2$ and $\sigma_{12i}(2) = 3$. Then $\sigma_{12i}(i = 1)$ is the formula $i \geq 2 \wedge i < 3$. In fact, for this embedding, the threshold 1 is translated into an interval $[2;2]$ containing the unique value 2. For the variable j, we have $\sigma_{12j}(1) = 1$ and $\sigma_{12j}(2) = 3$. Thus, $\sigma_{12j}(j = 1)$ is the formula $j \geq 1 \wedge j < 3$ that represents the interval $[1;2]$ containing the two values 1 and 2.*

4.2. Satisfaction

The validity of formulas is expressed via a binary relation denoted by $\models$ between models and BRN formulas. This relation is inductively defined on the structure of formulas. This then requires first to give a meaning of atomic formulas in $Atom(G)$ for a BRN signature G. For this, we need to extend G-models p by introducing the labelling function L in $ATS(G,p)$ which associates to each state s of S_G the following set of of atomic formulas of $Atom(G)$ considered as true at s:

$$L(s) = \{(i \sim l) \in Atom(G) \mid s(i) \sim l\}$$

When $ATS(G,p)$ is provided with a labelling function L, the resulting transition system is usually called a Kripke frame. Then, for a state formula φ, $p \models \varphi$ if for any state $s \in S_G$, (p,s) satisfies φ, denoted by $(p,s) \models \varphi$. In the same way, for a path formula π, $p \models \pi$ if for any path ρ of (S_G, T), (p,ρ) satisfies π, denoted by $(p,\rho) \models \pi$. Formally, the formula satisfaction is defined as follows:

Definition 14 (Formula satisfaction). *Let G be a BRN-signature and let p be a G-model. Let $s \in S_G$ be a state and let ρ be a path in $ATG(G,p)$. Both relations $(p,s) \models$ and $(p,\rho) \models$ on state formulas and path formulas, respectively, are inductively defined as follows:*

- $(p,s) \models at$ *iff $at \in L(s)$ for $at \in Atom(G)$;*
- $(p,s) \models \neg\varphi$ *iff $(p,s) \not\models \varphi$*
- $(p,s) \models \varphi \wedge \varphi'$ *iff $(p,s) \models \varphi$ and $(p,s) \models \varphi'$*
- $(p,s) \models \mathbf{E}\,\pi$ *iff there exists a path $\rho \in path(s)$ s.t. $(p,\rho) \models \pi$*
- $(p,\rho) \models \varphi$ *iff $(p,first(\rho)) \models \varphi$*
- $(p,\rho) \models \neg\pi$ *iff $(p,\rho) \not\models \pi$*
- $(p,\rho) \models \pi \wedge \pi'$ *iff $(p,\rho) \models \pi$ and $(p,\rho) \models \pi'$*
- $(p,\rho) \models \pi \mathbf{U} \pi'$ *iff there exists a path θ, suffix path of ρ ($\rho \leq \theta$), s.t. $(p,\theta) \models \pi'$ and s.t. for all suffix paths η verifying $\rho \leq \eta < \theta$, $(p,\eta) \models \pi$.*

Let us point out that the satisfaction relation is defined for finite and infinite paths as in [13]. We then deviate from the standard interpretation of CTL where only total relations and infinite paths are considered.

5. Formulas preservation results

In this section, based on the path preservation results given in Section 3, we prove two preservation results. In section 5.1, we show that along strict embedding all BRN-formulas are preserved, while in Section 5.2 we show that along monotonous embedding only a subset of BRN-formulas is preserved. This subset will contain all the BRN-formulas whose all negations are directly adjacent to propositional variables in $Atom(G)$.

5.1. Along strict embedding

Theorem 4. *Let $G \hookrightarrow G'$ be a strict signature embedding, let p' be a G'-model, and let p be the reduced G-model from p'. For every formula $\varphi \in Sen(G)$,*

$$p' \models \bar{\sigma}(\varphi) \quad \Longleftrightarrow \quad p \models \varphi$$

Proof. Let us define the mapping $\mathfrak{L}' : S_{G'} \to 2^{\bar{\sigma}(Atom(G))}$ defined by:

$$\mathfrak{L}'(s') = \{\bar{\sigma}(\psi) | \exists s \in S_G, \psi \in L(s) \text{ and } s' \in B(s)\}.$$

It is obvious to show that for every $s, s' \in S_{G'}$ such that $s \simeq s'$, we have $\mathfrak{L}'(s) = \mathfrak{L}'(s')$. Indeed, by the definition of $\mathfrak{L}'$, $\mathfrak{L}'(s')$ only depends on the class $B(s)$ of $\simeq$ containing s'. We can then define the mapping $\mathfrak{L}'_{/\simeq} : S'_{G/\simeq} \to 2^{\bar{\sigma}(Atom(G))}$ by: $\mathfrak{L}'_{/\simeq}([s']) = \mathfrak{L}'(s')$.

Moreover, by Theorem 3, for every $s' \simeq s''$ and for every transition $(s', r') \in ATS(G', p')$, we know there exists a path $(s'_0, s'_1, \dots, s'_n) \in path(s'')$ such that for every $i < n$, $s'_i \simeq s'$ and $s'_n \simeq r'$. This then proves that $\simeq$ is a divergence blind stuttering equivalence (dbs for short) according to the definition given in [13]. Let us first recall the definition of a dbs relation R on Kripke frame (S, T, L): a binary relation R on S is called a *dbs relation* if, and only if it is symmetric and

$$r \ R \ s \Longleftrightarrow \begin{cases} L(r) = L(s) \\ (r, r') \in T \Rightarrow & \exists \text{ a path } (s_0, s_1, \dots, s_n) \text{ in } (S, T), n \geq 0, (s_0 = s) \\ & \wedge (\forall i < n, r \ R \ s_i) \wedge r' \ R \ s_n \end{cases}$$

[13] gives the following result for any dbs relation $\simeq_{dbs}$:

Theorem 5 ([13])**.** *For a Kripke frame (S, T, L) and for any two states r and s of S, the following statements are equivalent:*

1. *$r \simeq_{dbs} s$*
2. *for every $CTL^\star$-X formula φ, $((S, T, L), r) \models \varphi$ iff $((S, T, L), s) \models \varphi$*

Therefore, by applying this theorem for the dbs relation defined on BRN states, we get that (p', s) and (p', r) are elementary equivalent (i.e. they satisfy the same set of CTL*-X formulas), we can deduce that $ATS(G', p')$ equipped with the

mapping $\mathcal{L}'$ and its quotient by $\simeq$ are elementary equivalent on the BRN-formulas built over $\bar{\sigma}(Atom(G))$, i.e.

$$\forall \varphi \in Sen(G), \qquad (S_{G'}, T', \mathcal{L}') \models \bar{\sigma}(\varphi) \iff (S_{G'}, T', \mathcal{L}')_{/\simeq} \models \bar{\sigma}(\varphi). \qquad (3)$$

Now, we can observe that the definition of B is traced on the renaming of atomic formulas in $Atom(G)$, i.e.

$$\forall s \in S_G, \ \forall s' \in B(s), \ \forall at \in Atom(G), (p, s) \models at \iff (p', s') \models \bar{\sigma}(at)$$

Then, following this observation and the definition of $\mathcal{L}'$, for every $s \in S_G$, $s' \in B(s)$ and $\varphi \in Atom(G)$, we have:

$$\begin{aligned}(p', s') \models \bar{\sigma}(\varphi) &\iff \varphi \in L(s) \\ &\iff \bar{\sigma}(\varphi) \in \mathcal{L}'(s') \\ &\iff \bar{\sigma}(\varphi) \in \mathcal{L}'_{/\simeq}(B(s)). \end{aligned} \qquad (4)$$

On one hand, it comes from this that,

$$\forall at \in Atom(G), \qquad p' \models \bar{\sigma}(at) \iff (S_{G'}, T', \mathcal{L}') \models \bar{\sigma}(at).$$

It is then easy to show, by induction on the structure of formulas, that

$$\forall \varphi \in Sen(G), \qquad p' \models \bar{\sigma}(\varphi) \iff (S_{G'}, T', \mathcal{L}') \models \bar{\sigma}(\varphi). \qquad (5)$$

On the other hand, it comes from (4) that,

$$\forall at \in Atom(G), \qquad (S_{G'}, T', \mathcal{L}')_{/\simeq} \models \bar{\sigma}(at) \iff p \models at.$$

Since, by Theorem 2 we know that transitions between different states in $ATS(G, p)$ and $ATS(G', p')_{/\simeq}$ are preserved, it is then easy to show, by induction on the structure of formulas, that

$$\forall \varphi \in Sen(G), \qquad (S_{G'}, T', \mathcal{L}')_{/\simeq} \models \bar{\sigma}(\varphi) \iff p \models \varphi. \qquad (6)$$

From (5), (3) and (6) we obtain:

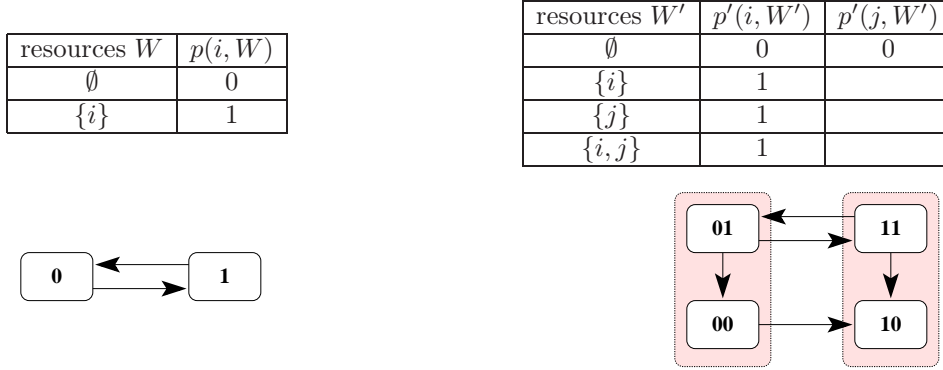$$\forall \varphi \in Sen(G), \qquad p' \models \bar{\sigma}(\varphi) \iff p \models \varphi.$$

$\square$

Given a set of temporal properties Γ over a BRN-signature G, $\Gamma^{\bullet} = \{\varphi | \forall \ p$ G-Model, $(\forall \psi \in \Gamma, p \models \psi) \Rightarrow p \models \varphi\}$.

Corollary 1. *For every strict embedding $G \hookrightarrow G'$, and every sets of BRN-formulas $\Gamma \subseteq Sen(G)$ and $\Gamma' \subseteq Sen(G')$ such that $\bar{\sigma}(\Gamma) \subseteq \Gamma'$, we have $\bar{\sigma}(\Gamma^{\bullet}) \subseteq \Gamma'^{\bullet}$.*

Proof. Let $\varphi \in \Gamma^{\bullet}$, let p' be a G'-Model such that for every $\psi \in \Gamma'$, $p' \models \psi$ and let p be the reduced G-model from p'. As $\bar{\sigma}(\Gamma) \subseteq \Gamma'$, by Theorem 4, we have $p \models \Gamma$. Therefore, $p \models \varphi$, and then by Theorem 4 we can conclude $p' \models \bar{\sigma}(\varphi)$. $\square$

FIGURE 10. An example of monotonous embedding.

resources W	$p(i, W)$
$\emptyset$	0
$\{i\}$	1

resources W'	$p'(i, W')$	$p'(j, W')$
$\emptyset$	0	0
$\{i\}$	1	
$\{j\}$	1	
$\{i, j\}$	1	

FIGURE 11. a G'-model p' (right) and its reduced G-model p (left).

5.2. Along monotonous embedding

Before giving our formula preservation result in the presence of monotonous embedding, we show by a very simple example that some BRN-formulas cannot be preserved along such embedding. Indeed, let us consider the two BRN-signatures G and G' of Figure 10. It is easy to see that G (left in the figure) is embedded into G' (right in the figure) and that this embedding is monotonous but not strict. Let p' be a G'-model and let p be the reduced G-model p from p' (see Figure 11), as well as the two asynchronous transition systems generated from p and p'. It is easy to see that the models p' and p cannot satisfy the same set of BRN-formulas. For example the formula $\mathbf{EF}(i = 0)$, which means that the system will finally pass through a state where $i = 0$, is satisfied by p but not by p'.

We are going notwithstanding to show a preservation result but on a restricted form of CTL*-X formulas. This gives rise to a subset of BRN-formulas, called nCTL*-X formulas, whose negations are directly adjacent to atoms in $Atom(G)$. This subset is defined as follows:

$$
\begin{aligned}
SFor ::= \;& Atom \mid \\
& \neg Atom \mid SFor \wedge SFor \mid SFor \vee SFor \\
& \mathbf{E} \; PFor \\
PFor ::= \;& SFor \mid \\
& PFor \wedge PFor \mid PFor \vee PFor \\
& PFor \; \mathbf{U} \; PFor
\end{aligned}
$$

Let us point out that the **F** operator is still usable but not the **G** operator, since $\mathbf{E}((at \vee \neg at) \mathbf{U} \varphi)$ is a nCTL*-X formula with at in $Atom(G)$ and φ a path formula, that can also written as $\mathbf{EF} \varphi$.

Lemma 5. *Let $G \hookrightarrow G'$ be a monotonous signature embedding, let p' be a G'-model, and let p be the reduced G-model from p'. Let $ATS(G', p') = (S_{G'}, T')$ and $ATS(G, p) = (S_G, T)$ the asynchronous transition systems associated respectively to p' and p. Let s be a state in S_G and $\rho \in path(s)$ be a path in $ATS(G, p)$. For every formula χ belonging to nCTL*-X formulas over $Atom(G)$, we have:*

(a) *If χ is a state formula, then $(p, s) \models \chi \Rightarrow \forall s' \in B'(s), (p', s') \models \bar{\sigma}(\chi)$*
(b) *If χ is a path formula, then $(p, \rho) \models \chi \Rightarrow \forall \rho' \in B'(\rho), (p', \rho') \models \bar{\sigma}(\chi)$*

*Proof.* First, we consider the case of the state formulas.

(1) $\chi = a$, with $a \in Atom(G)$: as $s' \in B'(s)$, we have $(p, s) \models \chi$ iff $(p', s') \models \bar{\sigma}(\chi)$.
(2) $\chi = \neg a$, with $a \in Atom(G)$: deduced from (1).
(3) $\chi = \varphi \wedge \varphi'$: the fact that $(p, s) \models \varphi \wedge \varphi' \Rightarrow \forall s' \in B'(s), (p', s') \models \bar{\sigma}(\varphi \wedge \varphi')$ follows since, by induction $(p, s) \models \varphi \Rightarrow \forall s' \in B'(s), (p', s') \models \bar{\sigma}(\varphi)$ and $(p, s) \models \varphi' \Rightarrow \forall s' \in B'(s), (p', s') \models \bar{\sigma}(\varphi')$, finally the fact comes from $\bar{\sigma}(\varphi) \wedge \bar{\sigma}(\varphi') = \bar{\sigma}(\varphi \wedge \varphi')$ and from the semantics of the $\wedge$ connector.
(4) $\chi = \varphi \vee \varphi$. This case can be handled as the previous case (3).
(5) $\chi = \mathbf{E} \pi$: suppose that $(p, s) \models \mathbf{E} \pi$. Then there exists a path $\rho \in path(s)$ in $ATS(G, p)$ such that $(p, \rho) \models \pi$. By induction, $\forall \rho' \in B'(\rho), (p', \rho') \models \bar{\sigma}(\pi)$. Since $B'(\rho)$ contains at least a path starting at any state from $B'(first(\rho))$, this implies $\forall s' \in B'(s)$, there exists a path $\rho' \in path(s')$ verifying $(p', \rho') \models \bar{\sigma}(\pi)$. By definition this implies $\forall s' \in B'(s), (p', s') \models \bar{\sigma}(\mathbf{E} \pi)$.

Next, we consider the case of the path formulas.

(6) $\chi = \varphi$: by definition, $(p, \rho) \models \chi$ iff $(p, s) \models \varphi$ for $s = first(\rho)$. By induction, we have $\forall s' \in B'(s), (p', s') \models \bar{\sigma}(\varphi)$. By definition this implies $\forall s' \in B'(s), \forall \rho' \in path(s'), (p', \rho') \models \bar{\sigma}(\varphi)$. Thus, $\forall \rho' \in B'(\rho), (p', \rho') \models \bar{\sigma}(\varphi)$.
(7) $\chi = \pi \Delta \pi'$ with $\Delta \in \{\wedge, \vee\}$. These cases can be handled in the same way than the case (3).
(8) $\chi = \pi \mathbf{U} \pi'$. Suppose that $(p, \rho) \models \pi \mathbf{U} \pi'$. Then there exists a path θ with $\rho \leq \theta$ such that $(p, \theta) \models \pi'$ and for all η verifying $\rho \leq \eta < \theta, (p, \eta) \models \pi$. There exists j such that $\theta = \rho^j$. Let ρ' be a path in $B'(\rho)$ provided with its adequate family $(k_i)_{i \in Index(\rho)}$. By induction, $(p', \rho'^{k_j}) \models \bar{\sigma}(\pi')$ and for all $k, 0 \leq k < k_j, (p', \rho'^k) \models \bar{\sigma}(\pi)$. Thus $\forall \rho' \in B'(\rho), (p', \rho') \models \bar{\sigma}(\pi) \mathbf{U} \bar{\sigma}(\pi')$ holds.

$\square$

By analogy of the definition of the sets $B'(\rho)$, we also define $B(\rho)$ the set of all paths ρ' such that for all indexes i of $Index(\rho)$, there exists an index k_i such that the family $(k_i)_{i \in Index(\rho)}$ is strictly increasing, and $\forall j \in Index(\rho'), (\exists i, k_i \leq j < k_{i+1} \Rightarrow s'_j \in B(s_i))$ or $((\forall i \in Index(\rho), k_i \leq j) \Rightarrow s'_j \in B(s_{max(Index(\rho))}))$. As by construction $B'(s) \subseteq B(s)$ and $B'(\rho) \subseteq B(\rho)$, we deduce the following theorem:
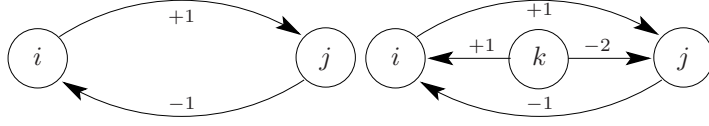
FIGURE 12. A non monotonous embedding $G \hookrightarrow G'$.

Theorem 6. *Let $G \hookrightarrow G'$ be a monotonous signature embedding, let p' be a G'-model, and let p be the reduced G-model from p'. Let $ATS(G', p') = (S_{G'}, T')$ and $ATS(G, p) = (S_G, T)$ the asynchronous transition systems associated respectively to p' and p. Let s be a state in S_G and $\rho \in path(s)$, with $s = first(\rho)$, be a path in $ATS(G, p)$. For every formula χ belonging to nCTL*-X formulas over $Atom(G)$, we have:*

(a) *If χ is a state formula, then $(p, s) \models \chi \Rightarrow \exists s' \in B(s), \ (p', s') \models \bar{\sigma}(\chi)$*
(b) *If χ is a path formula, then $(p, \rho) \models \chi \Rightarrow \exists \rho' \in B(\rho), \ (p', \rho') \models \bar{\sigma}(\chi)$*

The opposite implication is not true because there may exist some nCTL*-X formulas built over $Atom(G)$ that are satisfied by some state s' but not by s. For example we can see in Figure 8 that the formula $((0 \leq i < 2 \wedge j = 0) \Rightarrow \mathbf{EF} \ (j = 3))$ is satisfied by $((ATS(G', p'), (000))$ but $((i = 0 \wedge j = 0) \Rightarrow \mathbf{EF} \ (j = 2))$ is not satisfied by $(ATS(G, p), (00))$.

5.3. Counter-example justifying our notion of monotonous signature embedding

In this section we give a counter-example to show the relevance of monotony of Definition 3. Let us consider both BRN-signatures G and G' of Figure 12. We have an embedding $G \hookrightarrow G'$ which is not monotonous. Let p' be a G'-model given in Figure 13 (left) and let p be the reduced G-model built from p' and $G \hookrightarrow G'$ (see Figure 13-right), and let us consider the asynchronous transition systems generated from p and p'.
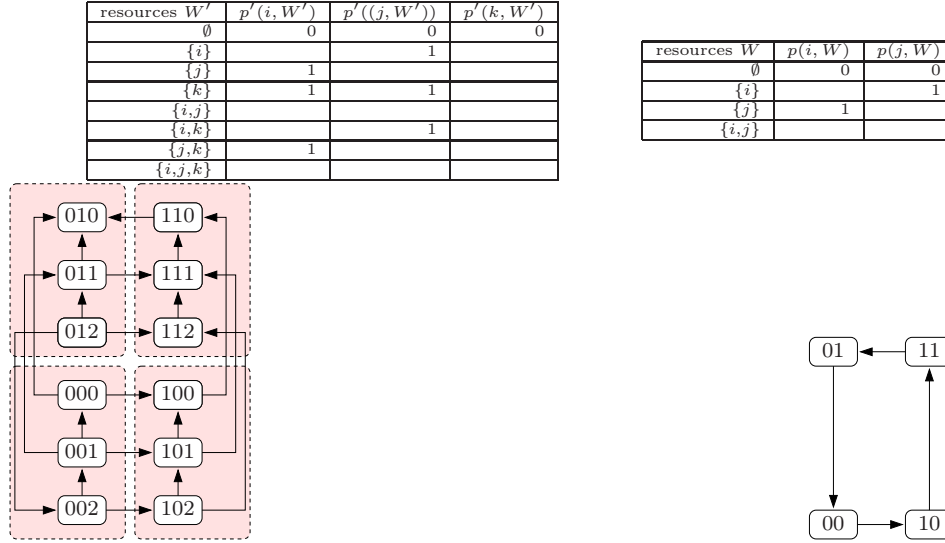It is then easy to see that Theorem 6 is not true. For example

$$(p, (11)) \models \varphi = ((i = 1 \wedge j = 1) \Rightarrow (\mathbf{EF}(i = 0 \wedge j = 0)))$$

while it does not exist $s' \in B((11))$ such that (p', s') satisfies $\bar{\sigma}(\varphi) = ((i = 1 \wedge j = 1) \Rightarrow (\mathbf{EF}(i = 0 \wedge j = 0)))$. That is:

$$\forall s' \in B((11)), (p', s') \not\models ((i = 1 \wedge j = 1) \Rightarrow (\mathbf{EF}(i = 0 \wedge j = 0)))$$

5.4. Discussion

The results of Sections 3 and 5.1 give us a modular way to study BRN dynamics. Indeed, let us consider a BRN-signature G'. We want to study the dynamics with respect to a set of basic biological experiment observations $\Gamma' \subseteq Sen(G')$. Moreover, let us suppose that G' is composed of a sub-BRN G whose dynamics have been already studied with respect to a set Γ such that $\bar{\sigma}(\Gamma) \subseteq \Gamma'$ is the subset of basic biological experiment observations about G. Hence, we suppose that the set $Dyn(G, \Gamma)$ of G-models satisfying the temporal properties in Γ, is known. Finally,

resources W'	$p'(i,W')$	$p'((j,W'))$	$p'(k,W')$
$\emptyset$	0	0	0
$\{i\}$		1	
$\{j\}$	1		
$\{k\}$	1	1	
$\{i,j\}$			
$\{i,k\}$		1	
$\{j,k\}$	1		
$\{i,j,k\}$			

resources W	$p(i,W)$	$p(j,W)$
$\emptyset$	0	0
$\{i\}$		1
$\{j\}$	1	
$\{i,j\}$		

FIGURE 13. A G'-model p' (left) and its reduced G-model p (right).

let us suppose that the embedding $G \hookrightarrow G'$ is strict. The procedure to construct $Dyn(G',\Gamma')$ is then the following:

For a generated G-model p in $Dyn(G,\Gamma)$, we generate all the G'-models whose reduced form is p. This step is easy: it consists in generating all the G'-models p' such that, for all $i \in V$ and $W \in G_i^-$,

$$p'(i,W) \in \{\sigma_i(p(i,W)), \ldots, \alpha_i(p(i,W))\}$$

where $\alpha_i(l) = \sigma_i(l+1)$ if $l < b_i$ and $\alpha_i(b_i) = b'_i$.

For each generated G'-model p', we apply model-checking techniques to check that p' satisfies temporal properties in $\Gamma' \setminus \bar{\sigma}(\Gamma)$.

With such a procedure, BRN dynamics can be studied step by step. Indeed, we start by studying the dynamics of all subgraphs G defining a strict embedding with G' and of minimal size (i.e. G does not contain other subgraphs also defining a strict embedding with G'). For each of them, we determine the maximal set Γ of biological experiment observations that rest on such subgraphs G and such that $\bar{\sigma}(\Gamma) \subseteq \Gamma'$. Then, we generate all the other minimal subgraphs G'' that contain one or more of the previous ones and are also strictly embedded in G'. We then compute their dynamics by using the previously described procedure. Finally, we restart this procedure until this is not possible any more. Computing minimal subgraphs defining a strict embedding is a computable problem having a polynomial complexity [9]. Finally, this procedure is obviously terminating because at each step we restrict the number of subgraphs defining a strict embedding with G'.

In practice, we have no assurance that the biological regulatory networks widely studied in the literature can be embedded in the whole biological regulatory network along an embedding that is strict or monotonous. Indeed, networks are often presented in the literature outside of any context. It is then difficult to know what is the potential influence of outside genes over the models of the studied models. Nevertheless, the work presented in this paper tends to question the validity of a biological property expressed as a formula that would be stated without knowing the nature of the embedding.

Considering only strict embeddings seems restrictive for practical applications, so the preservation of all CTL*-X properties is unlikely, although the discussion of this extreme case provides important intuition. Indeed, it was expected that strict embeddings give rise to strong preservation results. But it was not obvious at the beginning of the study that we would be able to identify another interesting class of embeddings, that is, the one of monotonous embeddings. Then, one can wonder how relevant the assumption of a monotonous embedding is for real biological systems. For instance, in the interaction graph between transcription factors in E. coli, that one can find in the data base RegulonDB [16], more than 90% of the interactions starting from a given vertex share, in average, the same sign. In fact, this high index is due to a large proportion of trivially monotonic regulators: those which are the source of a unique regulation with a unique sign. Nevertheless, if such regulators are excluded, the percentage goes down to 82% and thus remains rather high. The "monotonous assumption" then appears to be reasonable.

6. Related works

The work presented in this article is a continuation of earlier works made by some or all of the authors and can also be related to some other works.

6.1. Biological experiments as temporal formulas

As previously sketched, the main interest of modelling biological experiments as temporal formulas lays in the fact that it becomes possible to design dedicated automatic tools based on model-checking techniques. Applications of model checking techniques for the analysis and verification of qualitative models of biological regulatory networks have demonstrated their usefulness for understanding the dynamic behavior of these networks. The SMBioNet tool [5, 22] cuts down in the whole class of models to select the ones that satisfy some given biological experiment observations. In SMBioNet, biological experiment observations are expressed in Computation Tree Logic [14]. In [25], we propose an alternative approach combining symbolic execution and LTL model-checking techniques: instead of enumerating all the models satisfying a given (set of) formula(s), a constraint over the set of parameters $p(i, W)$ is computed such that any solution of the constraint provides a model satisfying the considered formulas.

We can also mention another tool whose purpose is very similar to the one of SMBioNet: the GNA tool [11] automatically checks that a given dynamic model satisfies some biological experiment observations expressed in CTL or in regular alternation-free μ-calculus [24, 4]. As in [25], [7] handles the family of parameters $p(i, W)$ under a symbolic form, using constraint logic programming to analyze properties of targeted biological regulatory networks.

These applications outlined certain limitations when dealing with large BRNs because of the classical state explosion problem. Our work aims at alleviating these limitations by promoting the usage of model checking on some small sub-BRNs. In the present paper, under some assumptions on the form of the embedding, we have obtained some property preservation results ensuring that formulas resting on the sub-BRN can be lifted at the level of the global BRN. Hence, every BRN satisfying such assumptions on some sub-BRN embeddings can be qualified as modular according to the meaning we gave to this word in [1, 2]. Indeed, we have introduced an abstract mathematical denotation highlighting the fact that a system may be qualified as complex (resp. modular) when some global properties cannot be (resp. are) directly derived from local properties satisfied by subsystems.

6.2. Preservation results along embedding of biological regulatory networks

This paper is a direct extension of [23] with complete proofs of the main results and with new results and additional examples. Among others, in [23] we proved that CTL-X formulas (i.e. the subset of state formulas in CTL*-X) are preserved along strict embedding of BRN. In this paper, we have weakened this condition by also considering monotonous embeddings, that allow us to preserve larger class of formulas. We have then shown that CTL*-X properties are preserved for strict embeddings, while we get a loose preservation result for monotonous embeddings.

In [6, 27], the authors also studied the preservation of BRN behaviors along embedding. In [6], the authors focus on dynamics associated to a BRN, and give a sufficient and necessary condition to strictly preserve the whole behavior (or dynamics) along an embedding. [27] takes place in a Boolean setting and gives some results on how to derive attractors for a global Boolean network from attractors of a subnetwork. The attractors in [27] are either steady states or sets of states where any trajectory issued from dynamics can not leave when it reaches one of the states constituting the attractor.

In this paper, we choose to study the preservation of temporal properties along embedding: a BRN embedding can preserve some temporal properties without preserving all (properties of) dynamics. This observational point of view (a temporal logic formula can be the transcription of a biological experiment observation) leads us to add hypotheses (strict or monotonous qualification) on embeddings to conclude on a property preservation.

7. Conclusion

In this paper, we have proposed to revisit the multi-valued discrete approach for biological regulatory networks using a logical formalism. BRN-signatures are made of graphs, denoting the static part of BRN. Formulas are temporal logic formulas over atoms expressing comparisons between concentration levels of gene products with some abstract discrete values. Models are asynchronous transition systems deduced from the knowledge of parametrization giving the concentration level toward which a variable is attracted. Lastly, the satisfaction relation is simply deduced from the one defined for $CTL^\star$. In order to study how properties expressed on a small BRN are preserved or not when it is embedded within a larger one, we have equipped our BRN formalism with signature embedding. Their main particularity is that they capture the fact that a concentration level or threshold relative to a network is converted into an interval of concentration levels. We have proved in this paper two results depending on the nature of the embedding: on one hand, when considering strict embedding, $CTL^\star$-X properties are preserved; whereas, when considering monotonous signature embedding, $nCTL^\star$-X properties of the small BRN are also preserved in the large.

In complex system analysis, such as biological regulatory networks, fixed points and cyclic behaviors play a crucial role. On one hand, homeostatic mechanisms allow an organism or a cell to maintain internal equilibrium in face of external variations. This leads to the presence of cycles in the state transition system. On the other hand, differentiation mechanisms have often a counterpart which is expressed in the state transition system in term of multi-stationarity. These two features are fundamental in systems biology since they allow one to explain a huge panel of biological functions. Thus, it becomes important to investigate what are the conditions to have on embeddings and/or models to preserve multi-stationarity and homeostasis. This last point will have to be connected to Siebert's works [27] who studies preservation of such properties but in Boolean network framework.

Acknowledgements

References

1. M. Aiguier, P. Le Gall, and M. Mabrouki, *A formal definition of complex software*, 3rd International Conference on Software Engineering Advances, IEEE Computer Society Press, 2008.

2. _____, *Complex software systems : Formalization and applications*, International Journal on Advances in Software **2** (2009), no. 1, 47–62.

3. G. Batt, C. Belta, and R. Weiss, *Temporal logic analysis of gene networks under parameter uncertainty*, IEEE Transactions on Automatic Control **53** (2008), no. z1, 215–29.

4. G. Batt, D. Ropers, H. De Jong, J. Geiselmann, R. Mateescu, M. Page, and D. Schneider, *Validation of qualitative models of genetic regulatory networks by model checking: Analysis of the nutritional stress response in escherichia coli*, Bioinformatics **21** (2005), no. Suppl 1, i19.

5. G. Bernot, J.P. Comet, A. Richard, and J. Guespin, *Application of formal methods to biological regulatory networks: Extending Thomas' asynchronous logical approach with temporal logic*, Journal of Theoretical Biology **229** (2004), no. 3, 339–47.

6. G. Bernot and F. Tahi, *Behaviour preservation of a biological regulatory network when embedded into a larger network*, Fundam. Inf. **91** (2009), no. 3-4, 463–85.

7. F. Corblin, E. Fanchon, and L. Trilling, *Applications of a formal approach to decipher discrete genetic networks*, BMC Bioinformatics **11** (2010), 385.

8. F. Corblin, S. Tripodi, E. Fanchon, D. Ropers, and L. Trilling, *A declarative constraint-based method for analyzing discrete genetic regulatory networks*, Biosystems **98** (2009), no. 2, 91–104.

9. T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein, *Introduction to algorithms*, 2nd ed., The MIT Press, 2001.

10. H. De Jong, *Modeling and simulation of genetic regulatory systems: a literature review.*, J. Comput. Biol. **9** (2002), no. 1, 67–103.

11. H. De Jong, J. Geiselmann, C. Hernandez, and M. Page, *Genetic network analyzer: qualitative simulation of genetic regulatory networks.*, Bioinformatics **19** (2003), no. 3, 336–44.

12. 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 **66** (2004), no. 2, 301–40.

13. R. De Nicola and F. Vaandrager, *Three logics for branching bisimulation*, J. ACM **42** (1995), no. 2, 458–87.

14. E.A. Emerson, *Handbook of theoretical computer science, volume b : formal models and semantics*, ch. Temporal and modal logic, pp. 995–1072, MIT Press, 1990.

15. F. Fages and S. Soliman, *Formal cell biology in biocham*, Formal Methods for Computational Systems Biology, 8th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, SFM 2008, Advanced Lectures, Lecture Notes in Computer Science, vol. 5016, Springer, 2008, pp. 54–80.

16. S. Gama-Castro, H. Salgado, M. Peralta-Gil, A. Santos-Zavaleta, L. Muniz-Rascado, H. Solano-Lira, V. Jimenez-Jacinto, V. Weiss, J.S. García-Sotelo, A. López-Fuentes, L. Porrón-Sotelo, S. Alquicira-Hernández, A. Medina-Rivera, I. Irma Martínez-Flores,

K. Alquicira-Hernández, R. Martínez-Adame, C. Bonavides-Martínez, J. Miranda-Ríos, A.M. Huerta, A. Mendoza-Vargas, L. Collado-Torres, B. Taboada, L. Vega-Alvarado, M. Olvera, L. Olvera, R. Grande, E. Morett, and J. Collado-Vides, *Regulondb version 7.0: transcriptional regulation of escherichia coli k-12 integrated within genetic sensory response units (gensor units).*, Nucleic Acids Res **39** (2011), no. Database issue, D98–105, http://regulondb.ccg.unam.mx/.

17. C.E. Giacomantonio, G.J. Goodhill, and K.J. Friston, *A boolean model of the gene regulatory network underlying mammalian cortical area development*, PLoS Computational Biology **6** (2010), no. 9, 251–62.

18. V.R. Glabbeck and W.P. Weijland, *Refinement in branching time semantics.*, Proc. IFIP Conference, 1989, pp. 613–18.

19. L. Glass and S.A. Kauffman, *The logical analysis of continuous non-linear biochemical control networks*, J. Theor. Biol. **39** (1973), 103–29.

20. J.A. Goguen and R.M. Burstall, *Institutions: Abstract model theory for specification and programming*, Journal of the ACM **39** (1992), no. 1, 95–146.

21. F. Jacob, , and J. Monod, *Genetic regulatory mechanisms in the synthesis of proteins*, Journal of Molecular Biology **3** (1961), 318–56.

22. Z. Khalis, J.P. Comet, A. Richard, and G. Bernot, *The SMBioNet method for discovering models of gene regulatory networks*, Genes, Genomes and Genomics **3(special issue 1)** (2009), 15–22.

23. M. Mabrouki, M. Aiguier, J.P. Comet, and P. Le Gall, *Property preservation along embedding of biological regulatory networks*, AB (K. Horimoto, G. Regensburger, M. Rosenkranz, and H. Yoshida, eds.), Lecture Notes in Computer Science, vol. 5147, Springer, 2008, pp. 125–38.

24. R. Mateescu and M. Sighireanu, *Efficient on-the-fly model checking for regular alternation-free mu-calculus*, Sci. Comput. Program. **46** (2003), 255–81.

25. D. Mateus, J.-P. Gallois, J.-P. Comet, and P. Le Gall, *Symbolic modeling of genetic regulatory networks*, Journal of Bioinformatics and Computational Biology **5** (2007), no. 2B, 627–40.

26. T. Schlitt and A. Brazma, *Current approaches to gene regulatory network modelling*, BMC bioinformatics **8** (2007), no. Suppl 6, S9.

27. H. Siebert, *Deriving behavior of boolean bioregulatory networks from subnetwork dynamics*, Math. Comput. Sci. **2** (2009), 421–42.

28. D. Thieffry, *From global expression data to gene networks*, BioEssays **21** (1999), no. 11, 895–99.

29. R. Thomas, *Logical analysis of systems comprising feedback loops.*, J. Theor. Biol. **73** (1978), no. 4, 631–56.

30. R. Thomas and R. d'Ari, *Biological feedback*, CRC Press, 1990.

31. H. Wehrheim, *Inheritance of temporal logic properties*, FMOODS, 2003, pp. 79–93.

Mbarka Mabrouki
Laboratoire d'informatique (LIX) F-91128 Palaiseau
e-mail: `mabrouki@lix.polytechnique.fr`

Marc Aiguier
Laboratoire de Mathématiques Appliquées aux Systèmes (MAS)
Grande Voie des Vignes - F-92295 Châtenay-Malabry
e-mail: marc.aiguier@ecp.fr

Jean-Paul Comet
Laboratoire I3S, UMR 6070 CNRS/UNSA
Algorithmes-Euclide-B, 2000 route des Lucioles
B.P. 121, F-06903 Sophia-Antipolis
e-mail: comet@unice.fr

Pascale Le Gall
Laboratoire de Mathématiques Appliquées aux Systèmes (MAS)
Grande Voie des Vignes - F-92295 Châtenay-Malabry
Programme d'Epigénomique
5 rue Henri Desbruères - F-91030 Evry
e-mail: pascale.legall@ecp.fr

Adrien Richard
Laboratoire I3S, UMR 6070 CNRS/UNSA
Algorithmes-Euclide-B, 2000 route des Lucioles
B.P. 121, F-06903 Sophia-Antipolis
e-mail: richard@unice.fr