



**HAL**  
open science

# Derivation of Qualitative Dynamical Models from Biochemical Networks

Wassim Abou-Jaoudé, Jérôme Feret, Denis Thieffry

► **To cite this version:**

Wassim Abou-Jaoudé, Jérôme Feret, Denis Thieffry. Derivation of Qualitative Dynamical Models from Biochemical Networks. Computational Methods in Systems Biology - CMSB 2015, Olivier Roux; Jérémie Bourdon, Sep 2015, Nantes, France. pp.195-207, 10.1007/978-3-319-23401-4\_17. hal-01199243

**HAL Id: hal-01199243**

**<https://inria.hal.science/hal-01199243v1>**

Submitted on 14 Feb 2025

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

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



Distributed under a Creative Commons Attribution 4.0 International License

# Derivation of Qualitative Dynamical Models from Biochemical Networks<sup>\*</sup>

Wassim Abou-Jaoudé<sup>1,2</sup>, Jérôme Feret<sup>1</sup>, and Denis Thieffry<sup>2</sup>

<sup>1</sup> DI-ENS (INRIA/ÉNS/CNRS/PSL<sup>\*</sup>), Paris, France

<sup>2</sup> IBENS (ÉNS/CNRS/Inserm/PSL<sup>\*</sup>), Paris, France

wassim@biologie.ens.fr, feret@ens.fr, thieffry@ens.fr

**Abstract.** As technological advances allow a better identification of cellular networks, more and more molecular data are produced allowing the construction of detailed molecular interaction maps. One strategy to get insights into the dynamical properties of such systems is to derive compact dynamical models from these maps, in order to ease the analysis of their dynamics. Starting from a case study, we present a methodology for the derivation of qualitative dynamical models from biochemical networks. Properties are formalised using abstract interpretation. We first abstract states and traces by quotienting the number of instances of chemical species by intervals. Since this abstraction is too coarse to reproduce the properties of interest, we refine it by introducing additional constraints. The resulting abstraction is able to identify the dynamical properties of interest in our case study.

## 1 Introduction

As technological advances allow a better identification of cellular networks, more and more molecular data are produced allowing the construction of detailed molecular interaction maps. These maps form large and complex intertwined biochemical networks, which dynamical functioning is very hard to decipher. One approach to unravel the dynamical properties of such systems relies on the derivation of qualitative dynamical models from these maps, in order to ease the analysis of their dynamics [?,?].

Automatic methods for such derivations still lack of convenient trade-off between efficiency and accuracy. Some abstractions consist only in partitioning the state space (as in the Boolean semantics of BIOCHAM [?]). These abstractions are usually too conservative and fail in detecting properties of interest. They

---

<sup>\*</sup> This material is based upon works sponsored by the “École normale supérieure” (ÉNS) under an incitative action and by the Defense Advanced Research Projects Agency (DARPA) and the U. S. Army Research Office under grant number W911NF-14-1-0367. The views, opinions, and/or findings contained in this article are those of the authors and should not be interpreted as representing the official views or policies, either expressed or implied, of the “École normale supérieure”, the Defense Advanced Research Projects Agency, or the U. S. Department of Defense.

have to be refined by integrating an approximated quantitative description of the dynamics of the model in each partition class, as done in tropical approximations [?] and piecewise affine systems [?]. Yet, the latter methods provide no explicit bounds for numerical errors (at best an asymptotic estimation of them).

Our motivation is twofold: not only we want to design an automatic tool to derive accurate logical models from reaction networks, but also we want to better understand the process of logical modelling and its underlying implicit assumptions. To achieve these goals, we use abstract interpretation [?]. Abstract interpretation is a mathematical framework to formally relate the behaviour of programs or models, seen at different levels of abstraction. It can be used not only to establish formal comparisons between abstraction techniques, but also to derive new abstractions of the behaviour of programs or models.

Our approach is the following. In Sect. ??, we formalise the behaviour of reaction networks by keeping the exact number of instances of chemical species. In Sect. ??, we propose an abstraction in which the number of instances is sampled within a finite set of intervals. In Sect. ??, we refine this abstraction by taking into account three kinds of properties: we deal with mass preservation invariants in Sect. ??; we detect when the number of instances of a given chemical species cannot cross its sampling intervals in Sect. ??; we enrich the description of the behaviour of the models with information about the reaction rates and take into account the separation between time-scales in Sect. ?. More details are provided in an extended version of this article [?].

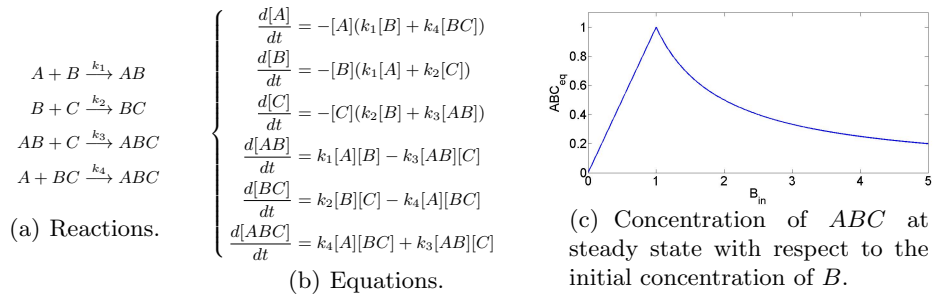
## 2 Case study

Let us start with a case study.

We consider a model with three kinds of proteins  $A$ ,  $B$ ,  $C$ . We assume that the protein  $B$  is a scaffold between the proteins  $A$  and  $C$ , that is to say that each instance of  $B$  can bind to an instance of  $A$  and/or to an instance of  $C$ . We wonder what is the influence of the initial concentration of the protein  $B$  on the concentration of the trimer  $ABC$ . Intuitively, the more  $B$ s we put in the model, the more  $ABC$ s will be formed. Yet this is not the case, since at high concentration, the protein  $B$  prevents the proteins  $A$  and  $C$  to meet since almost each instance of  $A$  (resp.  $C$ ) belongs to a dimer  $AB$  (resp.  $BC$ ), and thus there is no available  $A$ s (resp.  $C$ s) to form the trimers  $ABC$ . Thus, at high concentration, by sequestration effect, the scaffold prevents the formation of trimers  $ABC$ .

Fig.1 lists the reactions of the model (Fig. ??), the system of equations (under the assumptions of the law of mass action) (Fig. ??), and the concentration of the trimer  $ABC$  at steady state with respect to the initial concentration of the protein  $B$  (Fig. ??). We notice that at low concentration of the protein  $B$ , the concentration of the trimer  $ABC$  at steady state grows linearly, whereas it drops following an homographic function at high concentration of the protein  $B$ . Interestingly, this sequestration effect has also been observed *in vivo* [?].

This example is well suited for testing the accuracy of our approach, since two different dynamical behaviours may emerge according to the relative posi-



**Fig. 1.** A model with a sequestration effect. We plot the concentration of the trimer  $ABC$  at steady state with respect to the initial concentration of the protein  $B$ , with all reaction rates equal to 1 and with an initial concentration of 1 for the proteins  $A$  and  $C$  and of 0 for the complexes  $AB$ ,  $BC$ , and  $ABC$ .

tion of the quotient between the initial concentration of the protein  $B$  and the initial concentrations of the proteins  $A$  and  $C$  with respect to a semi-quantitative threshold. There is no need to know precisely the rates of the reactions. These quantitative details shift the threshold but have no impact on its existence (unless one of the reaction rate is set to 0). Although we have shown this phenomenon on the deterministic (differential) semantics, considering only forward reactions, it also occurs with a stochastic semantics and/or reversible reactions. A fine description of this model should account for complex properties such as concurrency and sequestration phenomena (when an instance of the protein  $A$  is bound to an instance of the protein  $B$ , it is no longer available to bind with an instance of the dimer  $BC$ ), as well as for the race between competing reactions (if there is many instances of  $B$  and few instances of  $BC$ , an instance of  $A$  is more likely to bind to a protein  $B$ , than to a complex  $BC$ ).

### 3 Trace semantics

We want to design a framework to automatically abstract logical models from reaction networks. Following a formal approach, we will relate the behaviour of the abstract model with the behaviour of the reaction network. Thus, the first task is to provide a formal definition for the behaviour of reaction networks. In this section, we describe this behaviour qualitatively in terms of a set of traces. Partial information about reaction kinetics will be taken into account in Sec. ??.

Firstly, we give the definition of a reaction network.

**Definition 1 (Reaction network).** *A network  $R$  of  $n$  reactions is a pair  $(\nu, (M_r, V_r)_{1 \leq r \leq n})$ , where: (i)  $\nu$  is a set of chemical species; (ii) for each integer  $r$  between 1 and  $n$ , (a)  $M_r : \nu \rightarrow \mathbb{N}$  is a multi-set of chemical species, and (b)  $V_r : \nu \rightarrow \mathbb{Z}$  is a reaction vector, such that  $M_r(x) + V_r(x) \geq 0$  for any  $x \in \nu$ .*

In Def. ??, a pair  $(M_r, V_r)$  is called a reaction. In a reaction  $(M_r, V_r)$ , the multi-set  $M_r$  encodes the set of the reactants (with their multiplicities) whereas the vector  $V_r$  denotes how many chemical species of each kind is produced and consumed at each application of the reaction.

We can now formally define the set of transitions of a reaction network.

**Definition 2 (Transition system).** A reaction network  $R \triangleq (\nu, (M_r, V_r)_{1 \leq r \leq n})$  induces a transition system  $(\mathcal{Q}_R, T_R)$  where: (i)  $\mathcal{Q}_R$  is the set  $\mathbb{N}^\nu$  of the functions between  $\nu$  and  $\mathbb{N}$ ; (ii)  $T_R$  is the subset of  $\mathbb{N}^\nu \times \llbracket 1, n \rrbracket \times \mathbb{N}^\nu$  that contains all the triple  $(q, r, q')$  such that, for all chemical species  $x \in \nu$ , (a)  $M_r(x) \leq q(x)$  and (b)  $q'(x) = q(x) + V_r(x)$ .

In Def. ??, the notation  $\llbracket 1, n \rrbracket$  denotes the set of the integers between 1 and  $n$ . The set  $\mathcal{Q}_R$  denotes all the potential states of the system. At this level of abstraction, the state of the system describes the number of instances of each kind of chemical species. The elements of  $T_R$  are called the transitions of the system. Transitions define the result of the applications of reactions. More precisely, a triple  $(q, r, q') \in T_R$  denotes the fact that the system can jump from the state  $q$  to the state  $q'$  by applying the rule indexed by the integer  $r$ . Condition (??) ensures that enough reactants are available, whereas condition (??) encodes the consumption/production of the chemical species. We notice that the resulting transition system is equivalent to a Petri net [?], in which each kind of chemical species is denoted by a placeholder and each instance by a token.

Before defining the traces of a reaction network, we introduce some notations. For any two sets  $A$  and  $\Sigma$ , and any subset  $T$  of the set  $A \times \Sigma \times A$ , we call a pretrace of elements of  $A$  and transitions in  $T$ , any element of the set  $A \times T^*$ . In a pretrace  $\tau \triangleq (a'_0, (a_i, \lambda_i, a'_i)_{1 \leq i \leq k})$ , the element  $a'_0$  (resp.  $a'_k$ ) is called the initial (resp. final) state of the pretrace  $\tau$  and is denoted as  $first(\tau)$  (resp.  $final(\tau)$ ). The second element of a pretrace is a (potentially empty) sequence of triples in  $T$ . We call a trace any pretrace  $(a'_0, (a_i, \lambda_i, a'_i)_{1 \leq i \leq k})$  such that  $a_i = a'_{i-1}$  for any integer  $i$  between 1 and  $k$ . Lastly, given a triple  $(a_{k+1}, \lambda_{k+1}, a'_{k+1})$  in  $T$ , we define by  $\tau \frown (a_{k+1}, \lambda_{k+1}, a'_{k+1})$  the pretrace  $(a'_0, (a_i, \lambda_i, a'_i)_{1 \leq i \leq k+1})$ .

We can now properly define the trace semantics of a reaction network.

**Definition 3 (Trace semantics).** The set of traces that is induced by a reaction network  $R$  and a set of initial states  $\mathcal{Q}_{R,0} \subseteq \mathcal{Q}_R$  is defined as the set of the traces  $\tau$  of elements of  $\mathcal{Q}_R$  and transitions in  $T_R$  such that  $first(\tau) \in \mathcal{Q}_{R,0}$ .

We denote by  $\mathcal{T}_{R, \mathcal{Q}_{R,0}}$  the set of traces that is induced by the reaction network  $R$  and the set of the initial states  $\mathcal{Q}_{R,0}$ .

Following the abstract interpretation framework [?], we can also express the trace semantics as the least fixpoint of a monotonic function over the powerset  $\wp(\mathcal{Q}_R \times T_R^*)$ . Let  $\mathbb{F}_{\mathcal{Q}_{R,0}}$  be the function that maps each set  $X$  of pretraces into the set of pretraces  $\mathcal{Q}_{R,0} \cup \{\tau \frown (q, r, q') \mid \tau \in X \wedge (q, r, q') \in T_R \wedge q = final(\tau)\}$ .  $\mathbb{F}_{\mathcal{Q}_{R,0}}$  is a monotonic function (i.e.  $X \subseteq Y \Rightarrow \mathbb{F}_{\mathcal{Q}_{R,0}}(X) \subseteq \mathbb{F}_{\mathcal{Q}_{R,0}}(Y)$ ) over a powerset, thus it has a least fixpoint [?]. This least fixpoint,  $lfp \mathbb{F}_{\mathcal{Q}_{R,0}}$ , is indeed

the set of all the traces of the reaction network  $R$  (i.e.  $\text{Lfp } \mathbb{F}_{\mathcal{Q}_{R,0}} = \mathcal{T}_{R, \mathcal{Q}_{R,0}}$ ). Moreover, the function  $\mathbb{F}_{\mathcal{Q}_{R,0}}$  is  $\cup$ -continuous (i.e.  $\mathbb{F}_{\mathcal{Q}_{R,0}}(\cup\{X_j \mid j \in J\}) = \cup\{\mathbb{F}_{\mathcal{Q}_{R,0}}(X_j) \mid j \in J\}$  for any family  $(X_j)_{j \in J}$  of sets of pretraces). It follows from [?] that the least fixpoint of  $\mathbb{F}_{\mathcal{Q}_{R,0}}$  can also be expressed as the limit of the finite iterates of the function  $\mathbb{F}_{\mathcal{Q}_{R,0}}$  (i.e.  $\mathcal{T}_{R, \mathcal{Q}_{R,0}} = \cup\{\mathbb{F}_{\mathcal{Q}_{R,0}}^i(\emptyset) \mid i \in \mathbb{N}\}$ ), which provides an iterative algorithm to enumerate the traces of the network  $R$ .

## 4 Derivation of a coarse-grained qualitative semantics

The semantics described in Sec. ?? is too fine grained. In particular, each instance of a protein is taken into account. Usually, in a qualitative model, the number of instances of proteins is sampled within a finite number of intervals. In this section, we will use the abstract interpretation framework to derive such an abstraction. Abstract interpretation [?] is a unifying framework for the approximation of mathematical structures. It offers formal tools to relate the observations of the behaviour of a system at different levels of details. It can also be used to systematically derive static analysers (that provide effective definitions of semantics at coarser levels of abstraction).

We use a simple version of the abstract interpretation framework that consists in removing some information from values, states and traces. Our abstraction is twofold. Firstly, we sample the number of instances of chemical species within a finite number of intervals. Secondly, we remove in traces the transitions for which the number of instances of each chemical species remain in the same interval. To sample the number of instances and later the rate of reactions (see Sect. ??), we partition the set  $\mathbb{R}^+$  over the  $p + 1$  intervals  $\llbracket 0, \delta \llbracket$ ,  $\llbracket \delta^i, \delta^{i+1} \llbracket$  for each integer  $i$  between 1 and  $p - 1$ , and  $\llbracket \delta^p, \infty \llbracket$ , where  $p$  and  $\delta$  are integer parameters such that  $\delta \geq 2$ . We introduce a function  $\beta^v$  to sample positive real numbers over this partition as follows:

**Definition 4 (abstract values).** *We define the function  $\beta^v$  between the set  $\mathbb{R}^+$  and the set  $\llbracket 0, p \llbracket$  that maps each positive real number  $v \in \mathbb{R}^+$  into the least integer in the set  $\{p\} \cup \{k \in \llbracket 0, p \llbracket \mid v < \delta^{k+1}\}$ .*

Then we lift the function  $\beta^v$  over transition systems.

**Definition 5 (abstract transition system).** *A reaction network  $R$  induces an abstract transition system  $(\mathcal{Q}_R^\sharp, T_R^\sharp)$  where: (i)  $\mathcal{Q}_R^\sharp$  is the set  $\llbracket 0, p \llbracket^\nu$  of the functions between the set of the chemical species  $\nu$  and the integer interval  $\llbracket 0, p \llbracket$ ; (ii)  $T_R^\sharp$  is the subset of  $\llbracket 0, p \llbracket^\nu \times \llbracket 1, n \llbracket \times \llbracket 0, p \llbracket^\nu$  that is defined by  $(q^\sharp, r, q'^\sharp) \in T_R^\sharp$  if and only if there exist  $(q, r, q') \in T_R$  such that  $q^\sharp = \beta^v \circ q$  and  $q'^\sharp = \beta^v \circ q'$ .*

Thus, the abstract transition system is obtained by applying component-wise the function  $\beta^v$  in the states of the transition system and in the states that occur in transitions. We denote by  $\beta^s$  the function mapping each state  $q \in \mathcal{Q}_R$  into the abstract state  $\beta^v \circ q \in \mathcal{Q}_R^\sharp$ . Then we lift the abstraction  $\beta^s$  to pretraces and traces. We call an abstract pretrace (resp. trace) any pretrace (resp. trace) of elements of  $\mathcal{Q}_R^\sharp$  and transitions in  $T_R^\sharp$ . We denote by  $\beta_1^\sharp$  the function between

the set  $\mathcal{Q}_R \times T_R^*$  and the set  $\mathcal{Q}_R^\# \times T_R^{\#*}$  that maps each (concrete) pretrace  $(q'_0, (q_i, r_i, q'_i)_{1 \leq i \leq k})$  to the (abstract) pretrace  $(\beta^s(q'_0), (\beta^s(q_i), r_i, \beta^s(q'_i))_{1 \leq i \leq k})$ . We notice that there exists some abstract transitions  $(q^\#, r, q^{\#'}) \in T_R^\#$  such that  $q^\# = q^{\#'}$ . Indeed, even if a concrete transition changes the number of instances of a chemical species, this does not always make it exit its sampling interval. We call such transitions silent and we denote by  $T_{R/\varepsilon}^\#$  the set of the non silent abstract transitions. In order to remove silent transitions, we define the function  $\beta_2^\#$  between the set  $\mathcal{Q}_R^\# \times T_R^{\#*}$  and the set  $\mathcal{Q}_R^\# \times T_{R/\varepsilon}^{\#*}$ , which maps each abstract pretrace  $(q_0^\#, (q_i^\#, r_i, q_i^{\#'}))$  to the abstract pretrace  $(q_0^\#, (q_{\sigma(i)}^\#, r_{\sigma(i)}, q_{\sigma(i)}^{\#'}))$ , where  $\sigma(i)$  ranges over the set  $\{i \in \llbracket 1, k \rrbracket \mid q_i^\# \neq q_i^{\#'}\}$  in increasing order.

We denote by  $\beta^t$  the composition of the function  $\beta_2^\#$  and  $\beta_1^\#$  and use this function to abstract the computation of the trace semantics. Given a set of initial states  $\mathcal{Q}_{R,0} \subseteq \mathcal{Q}_R$ , we introduce the function  $\mathbb{F}_{\mathcal{Q}_{R,0}}^\#$  over the set  $\wp(\mathcal{Q}_R^\# \times T_{R/\varepsilon}^{\#*})$  that is defined as  $\alpha^t \circ \mathbb{F}_{\mathcal{Q}_{R,0}} \circ \gamma^t$ , where: (i) the function  $\alpha^t$  maps each subset  $X$  of  $\mathcal{Q}_R \times T_R^*$  into the subset  $\{\beta^t(x) \in \mathcal{Q}_R^\# \times T_{R/\varepsilon}^{\#*} \mid x \in X\}$  of  $\mathcal{Q}_R^\# \times T_{R/\varepsilon}^{\#*}$  (ii) and conversely, the function  $\gamma^t$  maps each subset  $Y$  of  $\mathcal{Q}_R^\# \times T_{R/\varepsilon}^{\#*}$  into the subset  $\{x \in \mathcal{Q}_R \times T_R^* \mid \beta^t(x) \in Y\}$  of  $\mathcal{Q}_R \times T_R^*$ . Given two subsets  $X \subseteq \mathcal{Q}_R \times T_R^*$  and  $Y \subseteq \mathcal{Q}_R^\# \times T_{R/\varepsilon}^{\#*}$ , the property  $\alpha^t(X) \subseteq Y$  is equivalent to the property  $X \subseteq \gamma^t(Y)$ . Such a pair of functions is called a Galois connection [?]. Intuitively, the parts of the set  $\mathcal{Q}_R^\# \times T_{R/\varepsilon}^{\#*}$  denote properties about the elements in  $\mathcal{Q}_R \times T_R^*$ . The function  $\alpha^t$  abstracts each set of elements in  $\mathcal{Q}_R \times T_R^*$  into the most precise property they satisfy (the fact that  $(\alpha^t, \gamma^t)$  is a Galois connection entails that the most precise property always exists). Conversely, the function  $\gamma^t$  concretizes a property in  $\wp(\mathcal{Q}_R^\# \times T_{R/\varepsilon}^{\#*})$  into the set of the elements which satisfy this property. We define the Galois connections  $(\alpha^v, \gamma^v)$  (resp.  $(\alpha^s, \gamma^s)$ ) between sets of concrete values (resp. states) and sets of abstract values (resp. states) the same way.

The function  $\mathbb{F}_{\mathcal{Q}_{R,0}}^\#$  is monotonic. Thus, by [?], it has a least fixpoint.

**Definition 6 (abstract trace semantics).** *The set of abstract traces  $\mathcal{T}_{\mathcal{Q}_{R,0}}^\#$  that is induced by a reaction network  $R$  and a set of initial states  $\mathcal{Q}_{R,0} \subseteq \mathcal{Q}_R$  is defined as the least fixpoint  $\text{lfp } \mathbb{F}_{\mathcal{Q}_{R,0}}^\#$  of the function  $\mathbb{F}_{\mathcal{Q}_{R,0}}^\#$ .*

The Galois connection  $(\alpha^t, \beta^t)$  can be used to transfer the computation of the concrete fixpoint  $\mathcal{T}_{R, \mathcal{Q}_{R,0}} = \text{lfp } \mathbb{F}_{\mathcal{Q}_{R,0}}$  in the abstract.

**Theorem 1 (fixpoint transfer).** *For any reaction network  $R$  and any set of initial states  $\mathcal{Q}_{R,0} \subseteq \mathcal{Q}_R$ , the set  $\text{lfp } \mathbb{F}_{\mathcal{Q}_{R,0}}^\#$  is a subset of the set  $\gamma^t(\text{lfp } \mathbb{F}_{\mathcal{Q}_{R,0}})$ .*

We have used the Galois connection  $(\alpha^t, \gamma^t)$  so as to abstract the trace semantics. Thm ?? ensures that our abstraction is conservative, i.e. all the traces of the concrete semantics are taken into account. Moreover, the set of abstract traces can be computed by iterating the function  $\alpha^t \circ \mathbb{F}_{\mathcal{Q}_{R,0}} \circ \gamma^t$ . This consists in, at each step, (a) computing the concretization of the set of traces, (b) making the computation in the concrete, and (c) abstract the result.

The following property provides a direct way to make this computation without going back and forth in the concrete and provides more intuition about what

information is lost with our abstraction. We introduce few notations: we denote by  $V_\infty$  (resp.  $M_\infty$ ) the greatest element of the set  $\{|V_i(x)| \mid i \in \llbracket 1, n \rrbracket, x \in \nu\}$  (resp. of the set  $\{M_i(x) \mid i \in \llbracket 1, n \rrbracket, x \in \nu\}$ ); for any integer  $z \in \mathbb{Z}$ , we define the sign  $sign(z)$  of  $z$  as: (a)  $sign(0) \triangleq 0$ , and (b)  $sign(z) \triangleq z/|z|$  if  $z \neq 0$ ; and for any function  $f$  between two sets  $A$  and  $B$  and any elements  $y \in A$  and  $v \in B$ , we define  $f[y \mapsto v]$  as the function between  $A$  and  $B$  mapping the element  $y$  to the element  $v$ , and any element  $x \in A \setminus \{y\}$  to the element  $f(x)$ .

*Property 1.* The following assertions hold:

1. For any part  $Y \subseteq \mathcal{Q}_R^\# \times T_{R/\varepsilon}^{\#*}$ , the set  $\mathbb{F}_{\mathcal{Q}_{R,0}}^\#(Y)$  is equal to the set  $\alpha^s(\mathcal{Q}_{R,0}) \cup \{\tau^\# \frown (q^\#, r, q^{\#'}) \mid \tau^\# \in Y \wedge (q^\#, r, q^{\#'}) \in T_{R/\varepsilon}^\# \wedge final(\tau) = q^\#\}$ .
2. For any abstract transition  $(q^\#, r, q^{\#'}) \in T_{R/\varepsilon}^\#$ , if  $\delta > V_\infty$ , then the value  $q^{\#'}(x)$  is either equal to  $q^\#(x)$  or to  $q^\#(x) + sign(V_r(x))$ .
3. For any rule  $r$  and any abstract state  $q^\# \in \mathcal{Q}_R^\#$ , if  $\delta > M_\infty$ , then, for any chemical species  $y \in \nu$  such that  $V_r(y) \neq 0$  and  $0 \leq q^\#(y) + sign(V_r(y)) \leq p$ , we have  $(q^\#, r, q^\#[y \mapsto q^\#(y) + sign(V_r(y))]) \in T_R^\#$ .

Prop. ??(??) provides an inductive definition to compute the set of the abstract traces directly, without having to concretize the states. Prop. ??(??) establishes the fact that it is not possible to cross a whole interval in a single transition. As formalised in Thm. ??, the abstract trace semantics is a safe over-approximation of the concrete trace semantics. Yet, this semantics introduces spurious behaviours. In particular, Prop. ??(??) establishes that it is always possible to change the interval of a chemical species  $x \in \nu$  in the direction given by the sign of  $V_r(x)$ , when applying the rule that is indexed with the integer  $r$ , unless the chemical species  $x \in \nu$  is already in the first or in the last interval of the partition. This is a very coarse abstraction, which will be refined in the next section.

## 5 Refinements

As we have noticed in Sect. ??, the abstraction  $\mathcal{T}_{\mathcal{Q}_{R,0}}^\#$  is very coarse. In particular, it does not exploit the following three kinds of situations. Firstly, the number of instances of chemical species may be entangled by some mass preservation invariants. Secondly, when the number of instances of a chemical species enters a new interval, it is sometimes possible to prove that there are not enough resources in the system to make this number reach the next interval. Thirdly, our concrete semantics is purely qualitative. We propose to add kinetic rates and abstract them accurately in order to account for the potential races between reactions.

In this section, we propose three refinements of the abstract semantics to formalise three corresponding classes of reasoning. These refinements are orthogonal: they can be combined by the means of a reduced product [?].

### 5.1 Mass invariants

In the concrete semantics, the number of instances of the chemical species may be related by some mass conservation equations. For instance, in our case study,



the overall numbers of  $A$ s, of  $B$ s, and of  $C$ s are constant. Thus the expressions  $q(A) + q(AB) + q(ABC)$ ,  $q(B) + q(AB) + q(BC) + q(ABC)$ , and  $q(C) + q(BC) + q(ABC)$  keep the same values along a trace.

We are interested in constraints of the form  $\sum \alpha_x q(x) = b$  for  $(\alpha_x)_{x \in \nu} \in \mathbb{N}^\nu$  and  $b \in \mathbb{N}$  (i.e. semi-positive constraints). There are several ways to obtain the semi-positive constraints that are satisfied in a network. An algorithm that computes a basis of the set of the semi-positive relationships in a reaction network is proposed in [?]. Less costly but incomplete approaches can also be used: if we know the protein composition of chemical species, we can detect by scanning the set of the proteins and the set of the reactions which proteins are preserved by each reaction, and infer the corresponding mass preservation invariants.

Mass preservation invariants are particular cases of trace invariants and can thus be used to refine our abstraction. Let  $inv \subseteq \mathcal{Q}_R \times T_R^*$  be a trace invariant (formally, this means that  $\mathbb{F}_{\mathcal{Q}_{R,0}}(inv) \subseteq inv$ ). By [?], the concrete semantics is the most precise of the trace invariants (i.e.  $\mathcal{T}_{R,\mathcal{Q}_{R,0}} = \bigcap \{X \mid \mathbb{F}_{\mathcal{Q}_{R,0}}(X) \subseteq X\}$ ). In particular,  $\mathcal{T}_{R,\mathcal{Q}_{R,0}} \subseteq inv$ . It follows that  $\text{lfp } \mathbb{F}_{\mathcal{Q}_{R,0}} = \text{lfp } \mathbb{F}_{\mathcal{Q}_{R,0},inv}^{\text{INV}}$ , where  $\mathbb{F}_{\mathcal{Q}_{R,0},inv}^{\text{INV}}$  is defined as the function over the powerset  $\wp(\mathcal{Q}_R \times T_R^*)$  mapping each subset  $X \subseteq \mathcal{Q}_R \times T_R^*$  to the set  $\mathbb{F}_{\mathcal{Q}_{R,0}}(X) \cap inv$ . The least fixpoints of both functions  $\mathbb{F}_{\mathcal{Q}_{R,0}}$  and  $\mathbb{F}_{\mathcal{Q}_{R,0},inv}^{\text{INV}}$  are equal, but the abstraction of the iterates of the latter may be more precise. Let  $\mathbb{F}_{\mathcal{Q}_{R,0},inv}^{\text{INV}\#}$  be the function  $\alpha^t \circ \mathbb{F}_{\mathcal{Q}_{R,0},inv}^{\text{INV}} \circ \gamma^t$ . The function  $\alpha^t$  is  $\cap$ -complete, so the function  $\mathbb{F}_{\mathcal{Q}_{R,0},inv}^{\text{INV}\#}$  is equal to  $[Y \mapsto \mathbb{F}_{\mathcal{Q}_{R,0}}^\#(Y) \cap \alpha^t(inv)]$ . The iterates of the function  $\mathbb{F}_{\mathcal{Q}_{R,0},inv}^{\text{INV}\#}$  provide another effective way, more precise but still sound, to abstract the trace semantics:

**Theorem 2 (abstract trace semantics with invariants).** *Let  $\mathcal{Q}'_{R,0}$  be a subset of  $\mathcal{Q}_{R,0}$  and  $inv$  be a part of  $\mathcal{T}_{R,\mathcal{Q}_{R,0}}$  such that  $\mathbb{F}_{\mathcal{Q}'_{R,0}}(inv) \subseteq inv$ .*

*Then, we have:  $\mathcal{T}_{R,\mathcal{Q}'_{R,0}} \subseteq \gamma^t(\text{lfp } [Y \mapsto \mathbb{F}_{\mathcal{Q}'_{R,0}}^\#(Y) \cap \alpha^t(inv)])$ .*

In Thm. ??, we have partitioned the traces [?,?] to separate the computation of their abstraction according to their initial states. This leads to a more accurate abstraction whenever some pairs of initial states do not share the same invariants.

When the trace invariant is a set of semi-positive constraints, the following property gives an explicit definition for the term  $\alpha^t(inv)$ .

*Property 2 (mass invariant separation).* Let  $(a_x)_{x \in \nu} \in \mathbb{N}^\nu \setminus \{0\}^\nu$  be a family of positive integer coefficients (with at least one not equal to 0),  $b \in \mathbb{N}$  be a positive integer coefficient, and  $q^\#$  be an abstract state in  $\mathcal{Q}_R^\#$ . We denote by  $S$  the sum of the coefficients  $a_x$  for any chemical species  $x \in \nu$  and we introduce, for any abstract state  $q^\#$ ,  $q_{max}^\#$  as the maximum element of the set  $\{k \in \llbracket 0, p \rrbracket \mid \exists x \in \nu, a_x > 0 \wedge k = q^\#(x)\}$ . We further assume that  $S < \delta$ . Then, if either  $S\delta^{\beta^v(b)} \leq b$  or  $\beta^v(b) = 0$ , the set  $\alpha^s(\{q \in \mathcal{Q}_R \mid b = \sum_{x \in \nu} \alpha_x q(x)\})$  is equal to the set  $\{q^\# \in \mathcal{Q}_R^\# \mid q_{max}^\# = \beta^v(b)\}$ . Otherwise, it is equal to the set  $\{q^\# \in \mathcal{Q}_R^\# \mid q_{max}^\# \in \{\beta^v(b) - 1, \beta^v(b)\}\}$ .

Prop. ?? has a flavour of tropical algebrae [?]. In particular, whenever the affine constants of mass preservation invariants are far enough from the lower

bound of their sampling interval, the abstraction of the number of instances of a protein is equal to the abstraction of the number of instances of the most abundant chemical species containing this protein.

## 5.2 Watching interval boundaries

So far, we have approximated the number of instances of each chemical species by means of intervals. This is a quite coarse abstraction. Indeed, when the number of instances of a chemical species enters a new interval, there is no way to predict whether or not there may be enough resources in the system so that it may reach and enter the next interval. For instance, in our case study, when the system is in a state  $q \in \mathcal{Q}_R$  such that  $\beta^s(q)(A)$ ,  $\beta^s(q)(C)$  and  $\beta^s(q)(ABC)$  are all equal to 0, it may be possible to reach a state  $q'$  such that  $\beta^s(q')(ABC) = 1$ , because (1) the number of instances of  $ABC$  may be close to  $\delta$ , and (2) there may be enough instances of  $A$  to cross this threshold. But, after this, there will be not enough instances of  $A$  to reach a state  $q''$  such that  $\beta^s(q'')(ABC) > 1$ . We formalise this kind of reasoning and refine our abstraction accordingly.

We focus on proving that the number of instances of some chemical species cannot cross their current interval upwards (the dual case can be dealt with the same way). We assume that  $\delta > 2V_\infty$ . Given a state  $q \in \mathcal{Q}_R$  and a chemical species  $x \in \nu$ , we write  $q \models x_\dagger$  if either the value  $q(x)$  is in the interval  $[\delta^{\beta^v(q(x))}, \delta^{\beta^v(q(x))} + V_\infty]$  or if there is no concrete trace  $\tau$  in  $\mathcal{T}_{R, \{q\}}$  such that  $\beta^v(\text{final}(\tau))(x) > \beta^v(q(x))$ . We denote by  $\mathcal{C}_\dagger$  the set  $\{x_\dagger \mid x \in \nu\}$ .

We update the definitions of abstract states and abstract traces to take into account the constraints in  $\mathcal{C}_\dagger$ . Formally, an abstract state is now an element of  $\mathcal{Q}_R^\# \times \wp(\mathcal{C}_\dagger)$ . The first component encodes the intervals for the number of instances of chemical species, whereas the second component is a set of constraints that specifies which chemical species may eventually cross their current intervals upwards. We also define a refined abstraction function  $\beta_\dagger^s$  by  $\beta_\dagger^s(q) \triangleq (\beta^s(q), \{c \in \mathcal{C}_\dagger \mid q \models c\})$ . We denote by  $\beta_\dagger^t$  the function mapping each concrete trace  $\tau \in \mathcal{T}_{R, \mathcal{Q}_{R,0}}$  to the trace obtained by firstly replacing in the trace  $\tau$  every state  $q$  with its abstraction  $\beta_\dagger^s(q)$  and by secondly removing silent moves. The Galois connection that is induced by  $\beta_\dagger^s$  (resp.  $\beta_\dagger^t$ ) is denoted as  $(\alpha_\dagger^s, \gamma_\dagger^s)$  (resp.  $(\alpha_\dagger^t, \gamma_\dagger^t)$ ).

Iterating the most precise counterpart  $\alpha_\dagger^t \circ \mathbb{F}_{\mathcal{Q}_{R,0}} \circ \gamma_\dagger^t$  to the function  $\mathbb{F}_{\mathcal{Q}_{R,0}}$  would be very costly. Thus we iterate an over-approximation of it instead. We define  $esc$  as the set of the triples  $(q^\#, x_\dagger, r) \in \mathcal{Q}_R^\# \times \mathcal{C}_\dagger \times \llbracket 1, n \rrbracket$  such that that there is a concrete trace  $\tau \in \mathcal{T}_{R, \mathcal{Q}_R}$  which satisfies: (i)  $\beta^s(\text{first}(\tau)) = q^\#$ , (ii)  $\text{first}(\tau) \models x_\dagger$ , (iii)  $\beta^s(\text{first}(\tau))(x) < \beta^s(\text{final}(\tau))(x)$ , (iv)  $V_r(x) > 0$ , (v) there exists a transition in  $\tau$  with the label  $r$ . Intuitively, the set  $esc$  contains all the triples  $(q^\#, x_\dagger, r)$  such that, whenever the system is in a state  $q \in \gamma^s(\{q^\#\})$  satisfying  $q \models x_\dagger$ , the number of instances of the chemical species  $x$  may eventually cross the upper bound of its current interval, in a trace that contains at least one application of the rule indexed by the integer  $r$ . So as to offer a choice of trade-off between accuracy and efficiency, we introduce a superset  $esc^\#$  of  $esc$ , considered as a parameter of

our abstraction. Intuitively, whenever a triple  $(q^\sharp, x_\ddagger, r) \in esc^\sharp$ , it means that our approximation has failed in proving that the number of instances of the species  $x$  will never cross its current interval upwards.

We can now refine the set of the transitions of the abstract semantics.

**Definition 7 (abstract transitions).** We denote by  $T_{esc^\sharp}^{cross^\sharp}$  the set of the triples  $((q^\sharp, C), r, (q^{\sharp'}, C'))$  in  $(\mathcal{Q}_R^\sharp \times \wp(\mathcal{C}_\ddagger)) \times \llbracket 1, n \rrbracket \times (\mathcal{Q}_R^\sharp \times \wp(\mathcal{C}_\ddagger))$  such that:

1. either (a)  $(q^\sharp, r, q^{\sharp'}) \in T_R^\sharp$ , (b)  $\forall x_\ddagger \in C, q^{\sharp'}(x) \leq q^\sharp(x)$ , and (c)  $C' = (C \cup \{x_\ddagger \mid x \in \nu \wedge q^{\sharp'}(x) > q^\sharp(x)\}) \setminus \{x_\ddagger \mid x \in \nu \wedge q^{\sharp'}(x) < q^\sharp(x)\}$ ,
2. or (a)  $q^{\sharp'} = q^\sharp$  and there exists a constraint  $c \in \mathcal{C}_\ddagger$  such that: (b)  $C' = C \setminus \{c\}$ , and (c)  $(q^\sharp, c, r) \in esc^\sharp$ .

We distinguish between two kinds of transitions in Def. ???. The first ones consist in regular computation steps: they apply reactions that are allowed and do not violate the constraints about the capability of the chemical species to cross their intervals. After such reactions, the set of the chemical species that have just entered a new interval from below (resp. above) is recorded in (resp. removed from) the set of the constraints. The second kind of transitions consists in removing a constraint where we are unable to prove that the corresponding chemical species will never cross its current interval upwards.

Let  $\mathcal{T}_{esc^\sharp}^{cross^\sharp}$  be the set of pretraces of elements of  $\mathcal{Q}_R^\sharp \times \wp(\mathcal{C}_\ddagger)$  and transitions in  $T_{esc^\sharp}^{cross^\sharp}$ . Given a set of initial states  $\mathcal{Q}_{R,0} \subseteq \mathcal{Q}_R$ , we consider the function  $\mathbb{F}_{\mathcal{Q}_{R,0}, esc^\sharp}^{cross^\sharp}$  over the set  $\wp(\mathcal{T}_{esc^\sharp}^{cross^\sharp})$  mapping each subset  $Y$  of  $\mathcal{T}_{esc^\sharp}^{cross^\sharp}$  to the subset  $\alpha_\ddagger^s(\mathcal{Q}_{R,0}) \cup \{\tau^\sharp \frown (q^\sharp, r, q^{\sharp'}) \mid \tau^\sharp \in Y \wedge (q^\sharp, r, q^{\sharp'}) \in T_{esc^\sharp}^{cross^\sharp} \wedge final(\tau^\sharp) = q^\sharp\}$ . The function  $\mathbb{F}_{\mathcal{Q}_{R,0}, esc^\sharp}^{cross^\sharp}$  is monotonic and satisfies  $[\alpha_\ddagger^t \circ \mathbb{F}_{\mathcal{Q}_{R,0}} \circ \gamma_\ddagger^t](Y) \subseteq \mathbb{F}_{\mathcal{Q}_{R,0}, esc^\sharp}^{cross^\sharp}(Y)$  for any subset  $Y$  of  $\mathcal{T}_{esc^\sharp}^{cross^\sharp}$ . By [?], it follows that our approach is sound:

**Theorem 3 (soundness).** The function  $\mathbb{F}_{\mathcal{Q}_{R,0}, esc^\sharp}^{cross^\sharp}$  has a least fixpoint. Moreover, we have:  $lfp \mathbb{F}_{\mathcal{Q}_{R,0}} \subseteq \gamma_\ddagger^t(lfp \mathbb{F}_{\mathcal{Q}_{R,0}, esc^\sharp}^{cross^\sharp})$ .

The following property proposes a trade-off for the definition of the primitive  $esc^\sharp$ , based on a linear integer decision procedure.

*Property 3.* Let  $(q^\sharp, x_\ddagger, r) \in esc^\sharp$ . We have  $q^\sharp(x) \neq p$  and there exists a function  $w \in \mathbb{N}^{\llbracket 1, n \rrbracket}$  such that: (i)  $w(r) > 0$ , (ii)  $\delta^{q^\sharp(x)} + V_\infty + V_w(x) \geq \delta^{q^\sharp(x)+1}$ , and (iii)  $\forall x' \in \nu, q^\sharp(x') \neq p \Rightarrow \delta^{q^\sharp(x')+1} + V_w(x) > 0$ , where for any chemical species  $x' \in \nu$ ,  $V_w(x)$  denotes the value of the expression  $\sum_{1 \leq r' \leq n} w(r')V_{r'}(x')$ .

### 5.3 Scales separation

In our case study, when there are a lot of  $Bs$  and only a few  $BCs$  in the system, so as to capture the sequestration effect properly, we have to neglect the binding reaction between the chemical species  $A$  and  $BC$ . Thus we have to exploit the separation between different time scales. According to the modelling paradigm, several methods are used for the formalisation of the separation between time scales. In the logical approach, we usually assume that a reaction preempts any other much slower reactions. In the tropical approach, special care is taken not to

neglect the reactions which are involved in large time relaxations of fast cycles [?]. Another approach consists in encoding scales separation by the means of fairness hypotheses that bound the frequencies of slow reaction steps.

In this section, we propose (1) a generic method to formalise assumptions about time scale separation and (2) a systematic way to lift these assumptions to the abstract semantics. We start from a given scheduler  $\mathcal{S}$ . Formally,  $\mathcal{S}$  is a function between the set  $\mathcal{Q}_R \times T_R^*$  and the set  $\wp(\llbracket 1, n \rrbracket)$ . Intuitively, the scheduler restricts the set of the reactions which can be computed immediately after a (pre)trace. We refine the concrete semantics accordingly: we define  $\mathcal{T}_{\mathcal{R}, \mathcal{Q}_{R,0}, \mathcal{S}}^{\text{TIME}}$  as the least fixpoint of the monotonic function  $\mathbb{F}_{\mathcal{Q}_{R,0}, \mathcal{S}}^{\text{TIME}}$ , which maps any set  $X \subseteq \mathcal{Q}_R \times T_R^*$  of pretraces to the set  $\mathcal{Q}_{R,0} \cup \{\tau \frown (q, r, q') \mid \tau \frown (q, r, q') \in \mathbb{F}_{\mathcal{Q}_{R,0}}(X) \wedge r \in \mathcal{S}(\tau)\}$ .

Now we lift the action of the scheduler  $\mathcal{S}$  to the abstract semantics. For this end, we introduce, as a parameter of our analysis, a function  $\mathcal{S}^\sharp$  between the set  $\mathcal{Q}_R^\sharp \times T_{R/\varepsilon}^{\sharp*}$  and  $\wp(\llbracket 1, n \rrbracket)$  such that for any concrete trace  $\tau \in \mathcal{T}_{\mathcal{R}, \mathcal{Q}_{R,0}, \mathcal{S}}^{\text{TIME}}$  and any transition  $(q, r, q') \in T_R$  that satisfy (i)  $\text{final}(\tau) = q$ , (ii)  $\beta^s(q) \neq \beta^s(q')$ , and (iii)  $r \in \mathcal{S}(\tau)$ , we have  $r \in \mathcal{S}^\sharp(\beta^t(\tau))$ . Intuitively, a reaction  $r$  is in the set  $\mathcal{S}^\sharp(\tau^\sharp)$  whenever our approximation fails in proving that no trace  $\tau \in \gamma^t(\tau^\sharp)$  can be continued by applying the reaction  $r$  while changing the sampling interval of at least one chemical species. We introduce the function  $\mathbb{F}_{\mathcal{Q}_{R,0}, \mathcal{S}^\sharp}^{\text{TIME}^\sharp}$  over  $\wp(\mathcal{Q}_R^\sharp \times T_{R/\varepsilon}^{\sharp*})$  that maps each subset  $Y$  of  $\mathcal{Q}_R^\sharp \times T_{R/\varepsilon}^{\sharp*}$  to the subset  $\alpha^s(\mathcal{Q}_{R,0}) \cup \{\tau^\sharp \frown (q^\sharp, r, q^{\sharp'}) \mid \tau^\sharp \frown (q^\sharp, r, q^{\sharp'}) \in \mathbb{F}_{\mathcal{Q}_{R,0}}^\sharp(Y) \wedge r \in \mathcal{S}^\sharp(\tau^\sharp)\}$ . The function  $\mathbb{F}_{\mathcal{Q}_{R,0}, \mathcal{S}^\sharp}^{\text{TIME}^\sharp}$  is monotonic and satisfies  $[\alpha^t \circ \mathbb{F}_{\mathcal{Q}_{R,0}, \mathcal{S}}^{\text{TIME}} \circ \gamma^t](Y) \subseteq \mathbb{F}_{\mathcal{Q}_{R,0}, \mathcal{S}^\sharp}^{\text{TIME}^\sharp}(Y)$ , for any subset  $Y \subseteq \mathcal{Q}_R^\sharp \times T_{R/\varepsilon}^{\sharp*}$ . By [?], it follows that our approach is sound:

**Theorem 4 (soundness).** *The function  $\mathbb{F}_{\mathcal{Q}_{R,0}, \mathcal{S}^\sharp}^{\text{TIME}^\sharp}$  has a least fixpoint. Moreover, we have:  $\text{lfp } \mathbb{F}_{\mathcal{Q}_{R,0}, \mathcal{S}}^{\text{TIME}} \subseteq \gamma^t(\text{lfp } \mathbb{F}_{\mathcal{Q}_{R,0}, \mathcal{S}^\sharp}^{\text{TIME}^\sharp})$ .*

Let us instantiate our framework. For the sake of simplicity, we opt for the assumptions used in logical modelling, all the more so since there are no large time relaxation of fast cycles in our example. To each integer  $r \in \llbracket 1, n \rrbracket$ , we associate a kinetic function  $k_r$  between the set  $\mathcal{Q}_R$  and the set  $\wp(\mathbb{R}^+) \setminus \{\emptyset\}$ . The set  $k_r(q)$  denotes the potential propensity of the reaction indexed by  $r$  in the state  $q$  according to the (maybe partial) information that we may have about the rate of this reaction. The separation between time scales is encoded by a subset  $\text{Sep}$  of  $(\mathbb{R}^+)^2$  satisfying: (i) for any  $(x, y) \in \text{Sep}$ ,  $x < y$ ; (ii) for any  $x, y, x', y' \in \mathbb{R}^+$ , if  $(x, y) \in \text{Sep}$ ,  $x' \leq x$ , and  $y \leq y'$ , then  $(x', y') \in \text{Sep}$ . Intuitively, a pair  $(x, y)$  belongs to the set  $\text{Sep}$  when the value  $y$  is much greater than the value  $x$ . We define the concrete scheduler  $\mathcal{S}$  as the function mapping each pretrace  $\tau \in \mathcal{Q}_R \times T_R^*$  to the set of the reactions  $r$  such that for all reactions  $r'$ , we have  $k_r(\text{final}(\tau)) \times k_{r'}(\text{final}(\tau)) \not\subseteq \text{Sep}$ , meaning that the reaction  $r$  may be fast enough to exclude preemption by any other reaction.

In Prop. ??, we abstract away the dependency with respect to the concrete state  $\text{final}(\tau)$  so as to get an effective instantiation for the parameter  $\mathcal{S}^\sharp$ .

*Property 4.* Let  $(q^\sharp, r)$  be a pair in  $\mathcal{Q}_R^\sharp \times \llbracket 1, n \rrbracket$ . For any integer  $r' \in \llbracket 1, n \rrbracket$ , we denote by  $k_{r'}^\sharp(q^\sharp)$  the set of real values  $\alpha^v(\bigcup\{k_{r'}(q) \mid q \in \gamma^s(\{q^\sharp\})\})$ . If  $(\max(k_r^\sharp(q^\sharp)), \min(\bigcup\{k_{r'}^\sharp(q^\sharp) \mid r' \in \llbracket 1, n \rrbracket\})) \in \text{Sep}$ , then for any integer  $r' \in \llbracket 1, n \rrbracket$  and any state  $q \in \gamma^s(\{q^\sharp\})$ , we have  $(\max(k_r(q)), \min(k_{r'}(q))) \in \text{Sep}$ .

## 6 Conclusion

We have designed a formal framework to derive qualitative dynamical models from reaction networks. The results of the analysis of our case study is detailed in an extended version [?]. Interestingly, we can capture the sequestration effect: we can prove that when the number of instances of the protein  $B$  is very high (level 4) and those of the proteins  $A$  and  $C$  are low (level 2) in the initial state, then the number of instances of the complex  $ABC$  remains very low (levels 0, 1).

Our methodology offers a new trade-off between complexity and accuracy. It captures interesting properties that are beyond the scope of purely qualitative abstractions [?] and avoids the integration of numerical equations [?,?]. Our framework is purely formal and provides a better understanding of the qualitative modelling process, by clarifying the underlying assumptions. Interestingly, we notice that our approach often requires more intervals than in tropical approaches [?]. This is not so surprising, since in tropical approaches two consecutive intervals are assumed to be infinitely far from each other, whereas in our approach they contain arbitrarily close elements. One current limitation of our method is that we use one variable per chemical species, leading to a combinatorial explosion of the dynamics as the model size increases. To cope with this limitation, we plan to extend our framework to the reduced reaction networks obtained by the fragmentation [?,?] of the models written in the kappa language.

## References

1. <http://www.di.ens.fr/~feret/publications/CMSB2015/>
2. Bourdoncle, F.: Abstract interpretation by dynamic partitioning. *J. Funct. Prog.* 2(4) (1992)
3. Chaouiya, C.: Petri net modelling of biological networks. *Brief. Bioinfo.* 8(4) (2007)
4. Chapman, S.A., Asthagiri, A.R.: Quantitative effect of scaffold abundance on signal propagation. *Mol. Syst. Biol.* 5(1) (2009)
5. Cousot, P., Cousot, R.: Abstract interpretation: A unified lattice model for static analysis of programs by construction or approximation of fixpoints. In: *POPL* (1977)
6. Fages, F., Soliman, S.: Formal cell biology in biocham. In: *FMSB* (2008)
7. Feret, J., et al.: Internal coarse-graining of molecular systems. *PNAS* (2009)
8. Feret, J., et al.: Stochastic fragments: A framework for the exact reduction of the stochastic semantics of rule-based models. *IJSI* 7(4) (2013)
9. Grieco, L., et al.: Integrative modelling of the influence of MAPK network on cancer cell fate decision. *PLoS Comput. Biol.* 9(10) (2013)

10. de Jong, H., et al.: Qualitative simulation of genetic regulatory networks using piecewise-linear models. *Bull. Math. Biol.* 66(2) (2004)
11. Kleene, S.C.: *Introduction to mathematics*. ISHI Press International (1952)
12. Niarakis, A., et al.: Computational modeling of the main signaling pathways involved in mast cell activation. *Curr. Top. Microbio. Immun.* 382 (2014)
13. Radulescu, O., et al.: Reduction of dynamical biochemical reactions networks in computational biology. *Front. Genet.* 3(131) (2012)
14. Rival, X., Mauborgne, L.: The trace partitioning abstract domain. *TOPLAS* 29(5) (2007)
15. Schuster, S., Höfer, T.: Determining all extreme semi-positive conservation relations in chemical reaction systems: a test criterion for conservativity. *J. Chem. Soc.* 87 (1991)
16. Tarski, A.: A lattice-theoretical fixpoint theorem and its applications. *Pacific J. Math.* 5(2) (1955)