



**HAL**  
open science

# Graphical Conditions for Rate Independence in Chemical Reaction Networks

Elisabeth Degrand, François Fages, Sylvain Soliman

► **To cite this version:**

Elisabeth Degrand, François Fages, Sylvain Soliman. Graphical Conditions for Rate Independence in Chemical Reaction Networks. CMSB 2020 - 18th International Conference on Computational Methods in Systems Biology, Sep 2020, Konstanz / Online, Germany. <hal-02900758>

**HAL Id: hal-02900758**

**<https://inria.hal.science/hal-02900758v1>**

Submitted on 17 Jul 2020

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.



HAL Authorization

# Graphical Conditions for Rate Independence in Chemical Reaction Networks

Élisabeth Degrand, François Fages, and Sylvain Soliman

Inria Saclay-Île de France, Palaiseau, France

**Abstract.** Chemical Reaction Networks (CRNs) provide a useful abstraction of molecular interaction networks in which molecular structures as well as mass conservation principles are abstracted away to focus on the main dynamical properties of the network structure. In their interpretation by ordinary differential equations, we say that a CRN with distinguished input and output species computes a positive real function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , if for any initial concentration  $x$  of the input species, the concentration of the output molecular species stabilizes at concentration  $f(x)$ . The Turing-completeness of that notion of chemical analog computation has been established by proving that any computable real function can be computed by a CRN over a finite set of molecular species. Rate-independent CRNs form a restricted class of CRNs of high practical value since they enjoy a form of absolute robustness in the sense that the result is completely independent of the reaction rates and depends solely on the input concentrations. The functions computed by rate-independent CRNs have been characterized mathematically as the set of piecewise linear functions from input species. However, this does not provide a mean to decide whether a given CRN is rate-independent. In this paper, we provide graphical conditions on the Petri Net structure of a CRN which entail the rate-independence property either for all species or for some output species. We show that in the curated part of the Biomodels repository, among the 590 reaction models tested, 2 reaction graphs were found to satisfy our rate-independence conditions for all species, 94 for some output species, among which 29 for some non-trivial output species. Our graphical conditions are based on a non-standard use of the Petri net notions of place-invariants and siphons which are computed by constraint programming techniques for efficiency reasons.

## 1 Introduction

Chemical Reaction Networks (CRNs) are one fundamental formalism widely used in chemistry, biochemistry, and more recently computational systems biology and synthetic biology. CRNs provide an abstraction of molecular interaction networks in which molecular structures as well as mass conservation principles are abstracted away. They come with a hierarchy of dynamic Boolean, discrete, stochastic and differential interpretations [17] which is at the basis of a rich theory for the analysis of their qualitative dynamical properties [19,14,3], of their computational power [11,8,15], and on their relevance as a design method for

implementing high-level functions in synthetic biology, using either DNA [29,10] or DNA-free enzymatic reactions [12,32].

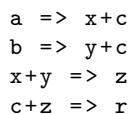
In their interpretation by ordinary differential equations, we say that a CRN with distinguished input and output species computes a positive real function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , if for any initial concentration  $x$  of the input species, the concentration of the output molecular species stabilizes at concentration  $f(x)$ . The Turing-completeness of that notion of chemical analog computation has been shown by proving that any computable real function can be computed by a CRN over a finite set of molecular species [15].

In the perspective of biochemical implementations with real enzymes however, the strong property of rate independence, i.e. independence of the computed result of the rates of the reactions [33], is a desirable property that greatly eases their concrete realization, and guarantees a form of absolute robustness of the CRN. The set of input/output functions computed by a rate-independent CRNs has been characterized mathematically in [9,4] as the set of piecewise linear functions. However, this does not give any mean to decide whether a given CRN is rate-independent or not.

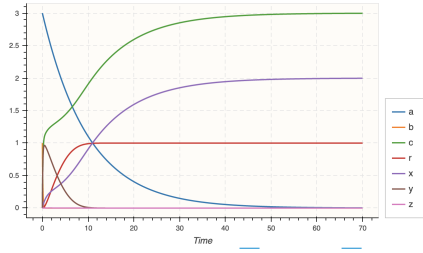
In this paper, we provide purely graphical conditions on the CRN structure which entail the rate-independence property either for all molecular species or for some output species. These conditions can be checked statically on the reaction hypergraph of the CRN, i.e. on its Petri net structure, or can be used as structural constraints in rate-independent CRN design problems.

*Example 1.* For instance, the reaction  $\mathbf{a+b} \Rightarrow \mathbf{c}$  computes at steady state the minimum of  $a$  and  $b$ , i.e.  $c^* = \min(a(0), b(0)) + c(0)$ ,  $a^* = \max(0, a(0) - b(0))$ ,  $b^* = \max(0, b(0) - a(0))$  whatever the reaction rate is. Our graphical condition for rate independence on all species assumes that there is no synthesis reaction, no fork and no loop in the reaction hypergraph (Thm. 2 below). This is trivially the case in this CRN and suffices to prove rate-independence for all species in this example.

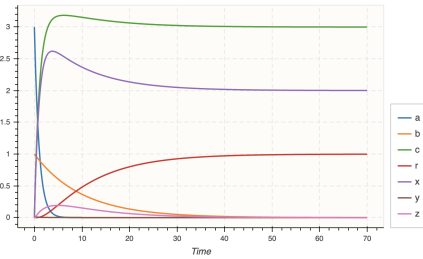
*Example 2.* Similarly, the CRN



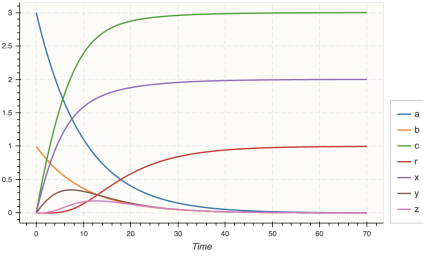
assuming  $x(0) = y(0) = c(0) = z(0) = r(0) = 0$ , computes at steady state the maximum of  $a$  and  $b$ :  $c^* = \max(a(0), b(0))$  (as  $a(0) + b(0) - \min(a(0), b(0))$ ),  $x^* = \max(0, a(0) - b(0))$ ,  $y^* = \max(0, b(0) - a(0))$ ,  $z^* = 0$ ,  $r^* = \min(a(0), b(0))$ ,  $a^* = 0$ ,  $b^* = 0$ , independently of the reaction rates. Fig. 1 shows some trajectories obtained with different values for the mass action law kinetics constants  $k_1, k_2, k_3, k_4$  of the four reactions above, with initial concentrations  $a(0) = 3, b(0) = 1$  and 0 for the other species. Here again, our graphical condition is trivially satisfied and demonstrates the rate-independence property of that CRN for all species, by Thm. 2.



(a)  $k_1 = 0.1, k_2 = 10.0, k_3 = 1, k_4 = 100.0$



(b)  $k_1 = 1, k_2 = 0.1, k_3 = 10, k_4 = 0.1$



(c)  $k_1 = 0.1, k_2 = 0.1, k_3 = 0.1, k_4 = 0.1$

Fig. 1: Computation of  $\max(a, b)$  with the rate-independent CRN of Ex. 2 with mass action law kinetics with different reaction rate constants.

The rest of this paper is organized as follows. In Sec. 3, we first give a sufficient condition for the rate independence of output species of the CRN. That condition tests the existence of particular P-invariants and siphons in the Petri net structure of the CRN. This test is modelled as a constraint satisfaction problem, and implemented using constraint programming techniques in order to avoid the enumeration of all P-invariants and siphons that can be in exponential number. Then in Sec. 4, we give another sufficient condition that entails the existence of a unique steady state, and ensures that the computed functions for all species of a CRN are rate-independent. None of these conditions are necessary conditions but we show with examples that they cover a large class of rate-independent CRNs. In Sec. 5, we evaluate our conditions on the curated part of the repository of models BioModels [7] by taking as output species the species

that are produced and not consumed. We show that 2 reaction graphs satisfy our rate-independence conditions for all species, 94 for some output species, among which 29 for some non-trivial output species. We conclude on the efficiency of our purely graphical conditions to test rate-independence of existing CRNs, and on the possibility to use those conditions as CRN design constraints for synthetic biology constructs such as [12].

## 2 Preliminaries

### 2.1 Notations

Unless explicitly noted, we will denote sets and multisets by capital letters (e.g.  $S$ , also using calligraphic letters for some sets), tuples of values by vectors (e.g.,  $\mathbf{x}$ ), and elements of those sets or vectors (e.g. real numbers, functions) by small Roman or Greek letters. For vectors that vary in time, the time will be denoted using a superscript notation like  $\mathbf{x}^t$ . For a multiset (or a set)  $M : S \rightarrow \mathbb{N}$ ,  $M(x)$  denotes the multiplicity of element  $x$  in  $M$  (usually the stoichiometry in the following), and 0 if the element does not belong to the multiset. By abuse of notation,  $\geq$  will denote the integer or Boolean pointwise order on vectors, multisets and sets (i.e. set inclusion), and  $+$ ,  $-$  the corresponding operations for adding or removing elements. With these unifying notations, set inclusion may thus be noted  $S \leq S'$  and set difference  $S - S'$ .

### 2.2 CRN Syntax

We recall here definitions from [16,18] for directed chemical reactions networks. In this paper, we assume a finite set  $S = \{x_1, \dots, x_n\}$  of molecular species.

**Definition 1.** A reaction over  $S$  is a triple  $(R, P, f)$ , where

- $R$  is a multiset of reactants in  $S$ ,
- $P$  a multiset of products in  $S$ ,
- and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a rate function over molecular concentrations or numbers.

A chemical reaction network (CRN)  $\mathcal{C}$  is a finite set of reactions.

It is worth noting that a molecular species in a reaction can be both a reactant and a product, i.e. a *catalyst*. Those mathematical definitions are mainly compatible with SBML([22]), however there are some differences. Unlike SBML, we find it useful to consider only directed reactions (reversible reactions being represented here by two reactions).

Furthermore, we enforce the following compatibility conditions between the rate function and the structure of a reaction:

**Definition 2 ([16,18]).** A reaction  $(R, P, f)$  over  $S$  is well-formed if the following conditions hold:

1.  $f$  is a non-negative partially differentiable function,
2.  $x_i \in R$  iff  $\partial f / \partial x_i(\mathbf{x}) > 0$  for some value  $\mathbf{x} \in \mathbb{R}_+^n$ ,
3.  $f(x_1, \dots, x_n) = 0$  iff there exists  $x_i \in R$  such that  $x_i = 0$ .

A CRN is well-formed if all its reactions are well-formed.

Those compatibility conditions are necessary to perform structural analyses of CRN dynamics. They ensure that the reactants contribute positively to the rate of the reaction at least in some region of the concentration space (condition 2), that the system remains positive (Prop. 2.8 in [16]) and that a reaction stops only when one of the reactant has been entirely consumed, whatever the rate function is.

To analyse the notion of function computed by a CRN, we will study the steady states of the ODE system, i.e., states where  $\frac{dx}{dt} = 0$ , the flux  $f_i$  of each reaction of  $\mathcal{C}$  at steady state will be called its *steady flux*.

A directed weighted bipartite graph  $\mathcal{G}_{\mathcal{C}}$  can be naturally associated to a chemical reaction network  $\mathcal{C}$ , with species and reactions as vertices, and stoichiometric coefficients, i.e. multiplicity in the multisets  $R$  and  $P$ , as weights for the incoming/outgoing edges.

*Example 3.* Fig. 2 shows the bipartite graph  $\mathcal{G}_{\mathcal{C}}$  of the Ex. 2 of the introduction. For this graph, the weights are all 1 and are not written for that reason.

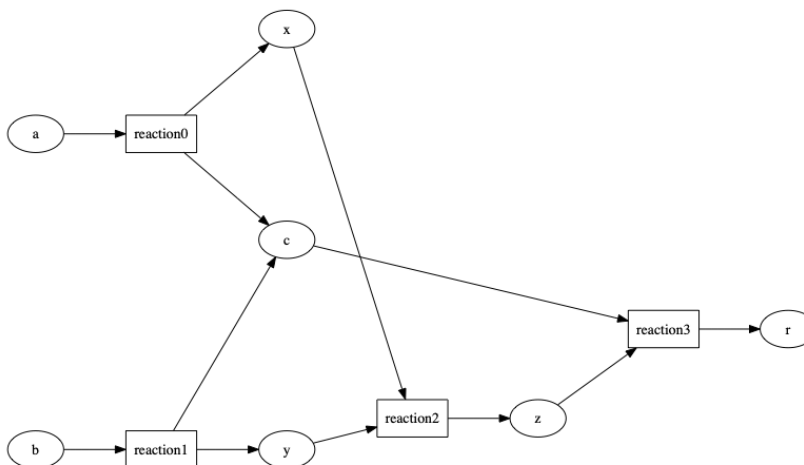


Fig. 2: Bipartite graph  $\mathcal{G}_{\mathcal{C}}$  associated to the CRN given in Example 2. The weights are all equal to 1 and not displayed.

### 2.3 CRN Semantics

As detailed in [17], a CRN can be interpreted in a hierarchy of semantics with different formalisms that can be formally related by abstraction relationships in the framework of abstract interpretation [13]. In this article, we consider the *differential semantics* which associates with a CRN  $\mathcal{C}$  the system  $\text{ODE}(\mathcal{C})$  of Ordinary Differential Equations

$$\frac{dx_j}{dt} = \sum_{(R_i, P_i, f_i) \in \mathcal{C}} (P_i(j) - R_i(j)) \cdot f_i$$

*Example 4.* Assuming mass action law kinetics for the CRN of Ex. 2, the ODEs are:

$$da/dt = -k_1 \cdot a \tag{1}$$

$$db/dt = -k_2 \cdot b \tag{2}$$

$$dc/dt = k_1 \cdot a + k_2 \cdot b - k_4 \cdot c \cdot z \tag{3}$$

$$dr/dt = k_4 \cdot c \cdot z \tag{4}$$

$$dx/dt = k_1 \cdot a - k_3 \cdot x \cdot y \tag{5}$$

$$dy/dt = k_2 \cdot b - k_3 \cdot x \cdot y \tag{6}$$

$$dz/dt = k_3 \cdot x \cdot y - k_4 \cdot c \cdot z \tag{7}$$

**Definition 3.** [15] *The function of time computed by a CRN  $\mathcal{C}$  from initial state  $\mathbf{x} \in \mathbb{R}^n$  is, if it exists, the solution of the ODE associated to  $\mathcal{C}$  with initial conditions  $\mathbf{x} \in \mathbb{R}^n$*

**Definition 4.** [15] *The input/output function computed by a CRN with  $n$  species, on an output species  $z$ , a set of  $m$  input species  $\mathbf{y} \in \mathbb{R}^m$  and a fixed initial state for the other species  $\mathbf{x} \in \mathbb{R}^{n-m}$  is, if it exists, the function  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  for which the ODEs associated to  $\mathcal{C}$  have a solution which moreover stabilizes on some value  $f(\mathbf{x}, \mathbf{y})$  on the  $z$  species component.*

**Definition 5.** *A CRN is rate-independent on an output species  $z$  if the input/output function computed on  $z$  with all species considered as input does not depend of the rate functions of the reactions.*

### 2.4 Petri Net Structure

The bipartite graph  $\mathcal{G}_{\mathcal{C}}$  of a CRN can be naturally seen as a Petri-net graph [30,5,6], here used with the continuous Petri-net semantics [21,20,31]. The species correspond to places and the reactions to transitions of the Petri-net. We recall here some classical Petri-net concepts [28,24] used in the next section, since they may have various names depending on the community.

**Definition 6.** A minimal semi-positive P-invariant is a vector of  $\mathbb{N}^n$  that is in the left-kernel of the stoichiometric matrix. Equivalently it is a weighted sum over places concentrations that remains constant by any transition.

A P-surinvariant is a weighted sum that only increases.

The support of a P-invariant or P-surinvariant is the set of places with non-zero value. Those places will be said to be covered by the P-invariant or P-surinvariant.

Intuitively a P-invariant is a conservation law of the CRN. The notion of P-surinvariant will be used to identify the output species of a CRN.

**Definition 7.** A siphon is a set of places such that for each edge from a transition to any place of the siphon, there is an edge from a place of the siphon to that transition.

Intuitively a siphon is a set of places that once empty remains empty, i.e., a set of species that cannot be produced again once they have been completely consumed. Our first condition for rate independence will be based on the following.

**Definition 8.** A critical siphon is a siphon that does not contain the support of any P-invariant.

A siphon that is not critical contains the support of a P-invariant, therefore it cannot ever get empty. A critical siphon on the other hand is thus a set of species that might disappear completely and then always remain absent.

### 3 Rate Independence Condition for Persistent Outputs

The persistence concept has been introduced to identify Petri nets for which places remain non-zero [1]. Here we establish a link between this notion of persistence and the rate-independence property of the input/output function computed on some output species.

#### 3.1 Sufficient Graphical Condition

As in [2], we are interested by the persistence not of the whole CRN but of some species. We will say that a species is an output of a CRN if it is produced and not consumed (and thus can only increase), i.e. if the stoichiometry of that species in the product part of any reaction is greater or equal to the reactant part, or equivalently:

**Definition 9.** A species is an output of a CRN if it is the singleton support of a P-surinvariant.

*Example 5.* In the CRN of Ex. 2, the *max* is computed on a non-output node, *c*. The *min* is computed on an output node *r*. The rate independence of that CRN on *r* follows from Thm. 1 below.

**Definition 10.** A species is structurally persistent if it is covered by a P-invariant and does not belong to any critical siphon.

Such species' concentrations will not reach zero for well-formed CRNs as proved in [1,2], but this section shows that if they are also output species they converge to a value that is independent of the rates of reactions. Note that such species might still belong to some non-critical siphons, for instance siphons that cover the whole P-invariant it is part of.

**Theorem 1.** If a species  $p$  of a well-formed CRN is a structurally persistent output, then that CRN is rate-independent on  $p$ .

*Proof.* Since  $p$  is structurally persistent, it is covered by some P-invariant and therefore bounded. Since  $p$  is an *output species* and the CRN is well-formed,  $\frac{dp}{dt} \geq 0$ . Hence its concentration converges to some value  $p^*$ .

When  $p$  reaches that steady state, all incoming reactions that modify it have null flux, hence by well-formed-ness one of their reactants has 0 concentration. If there are only incoming reactions that do not affect  $p$  then it is trivially constant and therefore rate-independent. Otherwise there are some such incoming reactions with a null reactant.

Now, notice that Prop. 1 of [1] states, albeit with completely different notations, that if one species of a well-formed CRN reaches 0 then all the species of a siphon reach 0. Therefore there exists a whole siphon  $S$  containing that reactant and with 0 concentration (intuitively, this reactant also has its input fluxes null, and one can thus build recursively a whole siphon).

By construction,  $S' = \{p\} \cup S$  is also a siphon, and since  $p$  is persistent,  $S'$  is not critical.  $S'$  therefore covers some P-invariant  $P$  and all concentrations are null except that of  $p$  in  $S'$ . Now  $P$  necessarily covers  $p$  since otherwise its conservation would be violated by having all 0 concentrations.

Note that by definition, for each P-invariant  $V$  containing  $p$  we have at any time  $t$  with state vector  $\mathbf{x}^t$  that  $V \cdot \mathbf{x}^t = V \cdot \mathbf{x}^0$ . Hence:

$$p(t) = \mathbf{x}^t(p) = \frac{V^0 - V \cdot \mathbf{x}_{p \rightarrow 0}^t}{V(p)} \leq \frac{V^0}{V(p)}$$

where  $\mathbf{x}_{p \rightarrow 0}^t$  is the state vector except for the concentration of  $p$  replaced by 0, and  $V^0$  is a shorthand for  $V \cdot \mathbf{x}^0$ .

At steady state we get  $p^* = \frac{P^0}{P(p)}$  since we proved that all concentrations other than that of  $p$  are null. Hence, we have:

$$p^* = \min_W \frac{V^0}{V(p)}$$

where  $W = \{V \mid V \text{ is a P-invariant covering } p\}$ , and which is obviously rate-independent.  $\square$

### 3.2 Constraint-based Programming

It is well-known that there may be an exponential number of P-invariants and siphons in a Petri net. Therefore, it is important to combine the constraints of both structural conditions for the computation of the minimal P-invariants and the union of critical siphons, without computing all siphons and P-invariants. This is the essence of constraint programming and of constraint-based modeling of such a decision problem as a constraint satisfaction problem. Furthermore, deciding the existence of a minimal siphon containing a given place is an NP-complete problem for which constraint programming has already shown its practical efficiency for enumerating all minimal siphons in BioModels, see [26].

We have thus developed a *constraint program* dedicated to the computation of structurally persistent species. For the minimal P-invariants, the constraint solving problem is the same as in [34] and is quite efficient on CRNs. For the second part about critical siphons, we use a similar approach but with Boolean variables to represent our siphons as in [26]. However, we enumerate *maximal* siphons here. This amounts to enumerate values 1 before 0, and to add in the branch-and-bound procedure for optimization that each new siphon must include at least one new place. Furthermore, we add the constraint that they are critical: for each P-invariant  $P$ , one of the species of its support must be absent (0). We get the flexibility of our constraint-based approach to add this kind of supplementary constraint while keeping some of the efficiency already demonstrated before.

In Section 5, this constraint program is used to compute the set of outputs and check if they are structurally persistent for many models of the `biomodels.net` repository. There are however a few models on which our constraint program is quite slow. An alternative constraint solving technique to solve those hard instances could be to use a SAT solver, at least for the enumeration of critical siphons, as shown in [26].

## 4 Global Rate Independence Condition

Thm 1 above can be used to prove the rate-independence property on some output species of a CRN, like  $r$  in Ex. 2 for computing the *max*, but not on some intermediate species, like  $c$  for computing the *min*. In this section we provide a sufficient condition for proving the rate-independence of a CRN on all species.

### 4.1 Sufficient Graphical Condition

**Definition 11.** *A chemical reaction network  $\mathcal{C}$  is synthesis-free if for all reactions  $(R_i, P_i, f_i)$  of  $\mathcal{C}$  we have  $R_i \not\leq P_i$ .*

In other words any reaction need to consume something to produce something.

**Definition 12.** *A chemical reaction network  $\mathcal{C}$  is loop-free if there is no circuit in its associated graph  $\mathcal{G}_{\mathcal{C}}$ .*

**Definition 13.** A chemical reaction network  $\mathcal{C}$  is fork-free if for all species  $x \in S$  there is at most one reaction  $(R_i, P_i, f_i)$  such that  $R_i(x) > 0$ .

This is equivalent to saying that the out-degree of species vertices is at most one in  $\mathcal{G}_{\mathcal{C}}$ .

**Definition 14.** A funnel CRN is a CRN that is:

1. synthesis-free
2. loop-free
3. fork-free

In Ex. 2 for computing the maximum concentration of two input species,  $A$  and  $B$ , one can easily check that the CRN satisfies the funnel condition (see Fig. 2). More generally, we can prove that any well-formed funnel CRN has a single stable state and that this state does not depend on the precise values of the parameters of the rate functions  $f_i$ .

**Lemma 1.** The structure of the bipartite graph  $\mathcal{G}_{\mathcal{C}}$  of a funnel CRN  $\mathcal{C}$  is a DAG with leaves that are only species.

*Proof.* Since  $\mathcal{C}$  is loop-free,  $\mathcal{G}_{\mathcal{C}}$  is acyclic. Since  $\mathcal{C}$  is synthesis-free, leaves cannot be reactions.  $\square$

**Lemma 2.** All steady fluxes of a funnel CRN  $\mathcal{C}$  are equal to 0.

*Proof.* Let us prove the lemma by induction on the topological order of reactions in  $\mathcal{G}_{\mathcal{C}}$ , this is enough thanks to Lemma 1.

For the base case (smallest reaction in the order), at least one of the species  $x$  such that  $R_i(x) > P_i(x)$  is a leaf (synthesis-freeness), then notice that at steady state,  $\frac{dx}{dt} = 0 = (P_i(x) - R_i(x))f_i$  since there is no production of  $x$  as it is a leaf, and no other consumption as  $\mathcal{C}$  is fork-free. Hence  $f_i = 0$ .

For the induction case, consider a reactant  $x$  s.t.  $R_i(x) > P_i(x)$  of our reaction. By induction hypothesis, at steady state we have  $\frac{dx}{dt} = 0 = (P_i(x) - R_i(x))f_i$  since all productions of  $x$  are lower in the topological order, and there is no other consumption of  $x$  as  $\mathcal{C}$  is fork-free. Hence  $f_i = 0$ .  $\square$

**Definition 15.** We shall denote  $x_i^+$  the total amount of species  $x_i$  available in an execution of the corresponding ODE system.

$$x_i^+ = x_i^0 + \int_0^{+\infty} \frac{dx_i^+}{dt} = x_i^0 + \int_0^{+\infty} \sum_{P_j(x_i) > R_j(x_i)} (P_j(x_i) - R_j(x_i))f_j$$

**Lemma 3.** Let  $\mathcal{C}$  be a well-formed funnel CRN, then for each initial state  $\mathbf{x}^0$ , if  $\mathcal{C}$  reaches a steady state  $\mathbf{x}^*$ , then the total amount  $x_i^+$  of any species  $x_i$  can be computed and is independent from the kinetic functions  $f_j$  of  $\mathcal{C}$ .

*Proof.* Let us proceed by induction on the topological order of species  $x_i$  in  $\mathcal{G}_C$ .

If  $x_i$  is a leaf, then since nothing produces it  $x_i^+ = x_i^0$ .

Now let us look at the induction case for  $x_i$ , and consider the set  $J$  of reactions producing  $x_i$  (i.e., such that  $P_j(x_i) > R_j(x_i)$ ).

From Lemma 2 we know that for all these reactions  $f_j = 0$  at stable state, and since  $\mathcal{C}$  is well-formed, it means that there exists at least one species  $x_{j_0}$  such that  $x_{j_0}^* = 0$ . As  $\mathcal{C}$  is fork-free and well-formed  $x_{j_0}$  has only been consumed by reaction  $r_j$ , which led to precisely producing an amount of  $x_i$  equal to  $x_{j_0}^+(P_j(x_i) - R_j(x_i))/(R_j(x_{j_0}) - P_j(x_{j_0}))$ , where  $x_{j_0}^+$  is available via induction hypothesis. Note also that  $j_0 = \operatorname{argmin}_{x_k | R_j(x_k) > P_j(x_k)} x_k^+(R_j(x_k) - P_j(x_k))$  since the reaction will stop as soon as it has depleted one of its inputs.

Hence  $x_i^+ = x_i^0 + \sum_J x_{j_0}^+(P_j(x_i) - R_j(x_i))/(R_j(x_{j_0}) - P_j(x_{j_0}))$ , which only depends on the initial state and the stoichiometry.  $\square$

**Theorem 2.** *Let  $\mathcal{C}$  be a well-formed funnel CRN, then the ODE system associated to  $\mathcal{C}$  has a single steady state  $\mathbf{x}^*$  that does not depend on the kinetic functions  $f_i$  of  $\mathcal{C}$ .*

*Proof.* From proof of Lemma 3 one notices that either  $x_i$  is not consumed at all and we have  $x_i^* = x_i^+$  or if  $j$  is the only reaction consuming  $x_i$ , its total consumption is given by  $x_{j_0}(R_j(x_{j_0}) - P_j(x_{j_0}))$ , with  $j_0$  defined as in the proof of Lemma 3.

These  $x_i^*$  do not depend on the kinetic functions  $f_i$  of  $\mathcal{C}$ .

We prove now that every  $x_i$  is convergent. It can be first noticed that

$$x_i(t) = x_{i_0} + F_e(t) - F_s(t)$$

where  $F_e(t) = \int_0^t \sum_{P_j(x_i) > R_j(x_i)} (P_j(x_i) - R_j(x_i)) f_j$  is the incoming flux and  $F_s(t) = \int_0^t (P_k(x_i) - R_k(x_i)) f_k$  is the outgoing flux.

$F_e$  is the integral of a positive quantity, it is then an increasing function. Moreover, as  $F_e(t) \leq x_i^+$ , this function is bounded and then converges to a real number limit.

Similarly,  $F_s$  is increasing and, as  $x_i(t) \geq 0$ , we have  $F_s(t) \leq x_i^+$  then it is bounded and converges.

To conclude,  $x_i$  is a difference of two convergent functions, hence it converges to a real number.

**Corollary 1.** *Any well-formed funnel CRN is rate-independent for any output species.*

We have thus given here a sufficient condition for a very strong notion of *rate-independence* in which all the species of the CRN have a steady state independent of the reaction rates, as in Ex. 2.

## 4.2 Necessary Condition

Our sufficient condition is not a necessary condition for global rate independence. Basically, forks that join and circuits that leak do not prevent rate independence:

*Example 6.* The CRN



is not a funnel CRN as it has both a loop (formed by  $a$  and  $b$ ) and a fork ( $b$  is a reactant in two distinct reactions). Nevertheless, this CRN is rate-independent on all species. The circuit formed by  $a$  and  $b$  has a leak with the third reaction. Every molecule of  $a$  and  $b$  will thus be finally transformed into  $c$  whatever the reaction kinetics are. At the steady state, the concentration of  $a$  and  $b$  will be null, and the concentration of  $c$  will be the sum of all the initial concentrations.

Nevertheless, we can show that any function computable by a rate independent CRN can be computed by a funnel CRN. We first show that funnel CRNs are composable under certain conditions for rate independent CRNs, similarly to the composability conditions given in [4].

**Definition 16.** *Two CRNs  $\mathcal{C}_1$  and  $\mathcal{C}_2$  are composable if*

$$\left( \bigcup_{(R,P,f) \in \mathcal{C}_1} R \cup P \right) \cap \left( \bigcup_{(R',P',f') \in \mathcal{C}_2} R' \cup P' \right) = \{x\}$$

*i.e., there is a single species appearing in both sets of reactions.*

*The composition of  $\mathcal{C}_1$  and  $\mathcal{C}_2$  is the union of their sets of reactions. The species  $x$  is called the link between both CRNs*

**Lemma 4.** *The composition of two funnel CRNs is a funnel CRN if the composition does not create forks on their link.*

*Proof.* As the reaction rates of the two original CRNs are well-formed, the reaction rates of the resultant CRN are well-formed too. No synthesis and no loop can be created by the union of two CRNs as all species are different except for the link. Therefore, the condition to create no fork by composition is sufficient to ensure that the resultant CRN is a funnel CRN.

**Corollary 2.** *The composition of two funnel CRNs is a funnel CRN if the link  $x$  is reactant in at most one of the CRNs.*

*Proof.* Since both CRNs are funnel,  $x$  is a reactant in at most one reaction in each. Now from our hypothesis it is not reactant at all in one of the CRNs, hence it appears as reactant in at most one reaction and therefore in no fork. By Lemma 4, the resulting CRN is a funnel CRN.

**Theorem 3.** *Any function computable by a rate independent CRN is computable by a funnel CRN.*

*Proof.* Using the same theorem from Ovchinnikov [27] as in [9] we note that any such function  $f$  with components  $f_j, 1 \leq j \leq p$  can be written as  $f(x) = \max_{1 \leq i \leq q} \min_{j \in S_i} f_j(x)$  for some family  $S_i \subseteq \{1, \dots, p\}$ .

Each  $f_j$  is rational linear, so this function can be written:  $f_j(x) = \sum_1^n \frac{\alpha_{j,i}}{n_j} x_i$ . To compute this linear sum, the following reactions are needed: For every  $x_i$ , we add the reactions  $x_i \Rightarrow \alpha_{j,i} \cdot w_j$  which compute  $w = \sum_1^n \frac{\alpha_{j,i}}{x} i$ .

Then we add the reaction  $n_j \cdot w_j \Rightarrow y_j$  which compute  $y_j = \frac{1}{n_j} w_j$ .

The output of the CRN that computes a linear function is a funnel CRN. Both max and min can be written with a funnel CRN (see respectively Example 2 and Example 2) and min can be composed by max as the output of min is not a reactant in the CRN that computes min. From Corollary 2, the conclusion is immediate.

## 5 Evaluation on Biomodels

In this section, we evaluate our sufficient condition for rate-independence on the reaction graphs of the curated part of the repository of models BioModels [7]. These models are numbered from BIOMD0000000001 to BIOMD0000000705. After excluding the empty models (i.e. models with no reactions or species), 590 models have been tested in total. As already noted in [16] however, many models in the curated of BioModels come from ODE models that have not been transcribed in SBML with well-formed reactions. Basically, some species appearing in the kinetics are missing as reactants or modifiers in the reactions, or some kinetics are negative. In this section, we test our graphical conditions for rate independence on the reaction graphs given for those models, without rewriting the structure of the reactions when they were not well-formed. Therefore, the actual rate-independence of the models that satisfy our sufficient criteria is conditioned to the well-formedness of the CRN.

The evaluation has been performed using Biocham<sup>1</sup> with a timeout of 240 seconds. The computer used for the evaluation has a quad-processor Intel Xeon 3.07GHz with 8Gb of RAM.

### 5.1 Computation of rate-independent output species

Following Def. 9, we tested the species that constitute the singleton support of a P-surinvariant. Among the 590 models tested, 340, i.e. 57.6% of them, were found to have no output species. 94 models, i.e. 15.9% of the models, were found to have at least one rate-independent output. 27 models, i.e. 4.5%, have both one rate-independent output and one undecided output, i.e. an output not satisfying our sufficient condition. 86 models, i.e. 14.5%, have at least one undecided output.

It is worth noting however that the species that are never modified by a reaction, i.e. that are only catalysts, remain always constant and thus constitute trivial rate-independent outputs. Amongst the 94 models with at least one rate-independent output found during evaluation, 29 have at least one non-trivial rate-independent output. Table 1 gives some details on the size and computation time for those 29 models.

<sup>1</sup> All our experiments are available on <https://lifeware.inria.fr/wiki/Main/Software#CMSB20b>

Biomodel#	#species	#reactions	#outputs	#RI	#NTRI	NTRI-species	Time (s)
037	12	12	2	2	2	Yi, Pi	0.950
104	6	2	3	3	1	species_4	0.074
105	39	94	11	3	1	AggP_Proteasome	63.366
143	20	20	4	1	1	MLTH_c	3.333
178	6	4	1	1	1	lytic	0.139
227	60	57	2	1	1	s194	17.299
259	17	29	1	1	1	s10	2.308
260	17	29	1	1	1	s10	2.310
261	17	29	1	1	1	s10	2.297
267	4	3	1	1	1	lytic	0.086
283	4	3	1	1	1	Q	0.053
293	136	316	14	4	3	aggE3, aggParkin, AggP_Proteasome	>240
313	16	16	4	2	1	IL13_DecoyR	2.071
336	18	26	1	1	1	IIa	4.148
344	54	80	7	2	1	AggP_Proteasome	>240
357	9	12	1	1	1	T	0.561
358	12	9	4	2	1	Xa_ATIII	0.892
363	4	4	1	1	1	IIa	0.067
366	12	9	4	2	1	Xa_ATIII	0.901
415	10	5	7	7	7	s10, s11, s12, s13, s14, s9, s15	0.894
437	61	40	22	8	1	T	16.109
464	14	10	6	3	1	s12	2.282
465	16	14	5	5	1	s23	59.554
525	18	19	8	3	1	p18inactive	33.479
526	18	19	8	3	1	p18inactive	33.858
540	22	11	12	11	8	s14, s15, s16, s17, s18, s19, s20, s21	56.134
541	37	32	13	9	7	s14, s15, s16, s17, s18, s19, s21	31.573
559	90	136	18	2	2	s493, s502	150.954
575	76	58	9	1	1	DA_GSH	66.806

Table 1: Model numbers in Biomodels containing non-trivial structurally persistent output species which are thus rate-independent by Thm. 1. For each model, we indicate the numbers of species, reactions, rate-independent species, non-trivial rate-independent species and total computation time in seconds.

Now, evaluating by simulation the actual rate-independence property of those models, and thereby the empirical completeness of our purely graphical criterion in this benchmark, would raise a number of difficulties. First, as said above, many SBML models coming from ODE models have not been properly transcribed with well-formed reactions and would need to be rewritten [16]. Second, some models may contain additional events or assignment rules which are not reflected in the

CRN reaction graph. Third, the relevant time horizon to consider for simulation is not specified in the SBML file. In the curated part of BioModels, this time horizon can range from 20s to 1 000 000s.

Nevertheless, we performed some manual testing on 9 models from Table 1, namely models 37, 104, 105, 143, 178 and 227, which have at least one non-trivial rate-independent output, and models 50, 52 and 54, which have only undecided outputs. For each model, numerical simulations were done with two different sets of initial concentrations and two different sets of parameters. Even when it was not the case in the original models, all the parameters were set to positive values. All outputs in models 37 and 104 were found rate-independent which was confirmed by numerical simulation. For model 105, 3 outputs among the 11 outputs of this model were found rate-independent by our algorithm which seemed again to be confirmed by numerical simulation. Models 143 and 227 are not well-formed which explains why the species satisfying our graphical criterion were shown not to be rate-independent by numerical simulation. For models with only undecided outputs, i.e. models 50, 52 and 54, numerical simulations show that none of their outputs is rate-independent. For these 3 models, 11 undecided outputs were tested in total. In this manual testing, we did not find any output that was left undecided by the algorithm and was found rate-independent by numerical simulation.

## 5.2 Test of global rate-independence

In this section, we test the criterion given in Def. 14 that ensures the rate-independence of all the species of a given CRN.

On the 590 reaction models tested, 20 models have reached the timeout limit of 240 seconds and were therefore not evaluated. Two models were found to be rate-independent on all species, namely models `BIOMD0000000178` and `BIOMD0000000267`. These models constitute a chain of respectively 4 and 3 species. At steady state, all species have a null concentration, except the last one. The steady state value of the last species is equal to the sum of all the initial concentrations. These models simulate the onset of paralysis of skeletal muscles induced by botulinum neurotoxin serotype A. They are used in particular to get an upper time limit for inhibitors to have an effect [25].

These two models were also found to have rate-independent outputs during the evaluation of the previous criterion for outputs. The global criterion here shows that not only the output species of the chain are rate-independent, but also all the inner species of the chain.

## 6 Conclusion

We have given two graphical conditions for verifying the rate-independence property of a chemical reaction network. First, the absence of synthesis, circuit and fork in the reaction graph, ensures the existence of a single steady state that does not depend on the reaction rates, thereby ensuring the existence of a

computed input/output function for all species of the CRN and their independence of the rate of the reactions. Second, the covering of a given output species by one P-invariant and no critical siphon, provides a criterion to ensure the rate-independence property of the computed function on that output species.

These graphical conditions are sufficient but none of them is necessary. Evaluation in BioModels suggests however that they are already quite powerful since among the 590 models of the curated part of BioModels tested, 94 reaction graphs were found rate-independent for some output species, 29 for non-trivial output species, and 2 for all species which was confirmed for well-formed models.

It is worth noting that our second condition uses the classical Petri net notions of P-invariant and siphons in a non-standard way for continuous systems. A similar use has already been done for instance in [1] for the study of persistence and monotone systems, and interestingly in [23], where the authors remark the discrepancy there is on the Petri net property of trap between the standard discrete interpretation, under which a non empty trap remains non empty, and the continuous interpretation under which a non empty trap may become empty. This shows the remarkable power of Petri net notions and tools for the study of continuous dynamical systems, thus beyond standard discrete Petri nets and outside Petri net theory properly speaking.

As already remarked in previous work [26,34], modeling the computation of Petri net invariants, siphons and other structural properties as a constraint satisfaction problem provides efficient implementations using general purpose constraint solvers, often showing better efficiency than with dedicated algorithms. This was illustrated here by the use of a constraint logic program to implement our condition on P-invariants and critical siphons by constraining the search to those sets of places that satisfy the condition, without having to actually compute the sets of all P-invariants and critical siphons.

Finally, it is also worth noting that beyond verifying the rate-independence property of a CRN and identifying the output species for which the computed function is rate-independent, our graphical conditions may also be considered as structural constraints to satisfy for the design of rate-independent CRNs in synthetic biology [12]. They should thus play an important role in CRN design systems in the future.

## Acknowledgements

This work was jointly supported by ANR-MOST *BIOPSY Biochemical Programming System* grant ANR-16-CE18-0029 and ANR-DFG *SYMBIONT Symbolic Methods for Biological Networks* grant ANR-17-CE40-0036.

## References

1. Angeli, D., Leenheer, P.D., Sontag, E.D.: A petri net approach to persistence analysis in chemical reaction networks. In: *Biology and Control Theory: Current Challenges*. LNCIS, vol. 357, pp. 181–216. Springer-Verlag (2007)

2. Angeli, D., Leenheer, P.D., Sontag, E.D.: Persistence results for chemical reaction networks with time-dependent kinetics and no global conservation laws. In: Proceedings of the 48th IEEE Conference on Decision and Control (CDC). pp. 4559–4564. IEEE (Dec 2009)
3. Baudier, A., Fages, F., Soliman, S.: Graphical requirements for multistationarity in reaction networks and their verification in biomodels. *Journal of Theoretical Biology* 459, 79–89 (Dec 2018), <https://hal.archives-ouvertes.fr/hal-01879735>
4. Chalk, C., Kornerup, N., Reeves, W., Soloveichik, D.: Composable rate-independent computation in continuous chemical reaction network. In: Češka, M., Šafránek, D. (eds.) *Computational Methods in Systems Biology*. pp. 256–273. Springer International Publishing (2018)
5. Chaouiya, C.: Petri net modelling of biological networks. *Briefings in Bioinformatics* 8(4), 210–219 (2007)
6. Chaouiya, C., Remy, E., Thieffry, D.: Petri net modelling of biological regulatory networks. *Journal of Discrete Algorithms* 6(2), 165–177 (Jun 2008)
7. Chelliah, V., Laibe, C., Novère, N.: Biomodels database: A repository of mathematical models of biological processes. In: Schneider, M.V. (ed.) *In Silico Systems Biology, Methods in Molecular Biology*, vol. 1021, pp. 189–199. Humana Press (2013)
8. Chen, H.L., Doty, D., Soloveichik, D.: Deterministic function computation with chemical reaction networks. *Natural computing* 7433, 25–42 (2012)
9. Chen, H.L., Doty, D., Soloveichik, D.: Rate-independent computation in continuous chemical reaction networks. In: Proceedings of the 5th Conference on Innovations in Theoretical Computer Science. pp. 313–326. ITCS '14, ACM, New York, NY, USA (2014)
10. Chen, Y., Dalchau, N., Srinivas, N., Phillips, A., Cardelli, L., Soloveichik, D., Seelig, G.: Programmable chemical controllers made from DNA. *Nature Nanotechnology* 8, 755–762 (Sep 2013)
11. Cook, M., Soloveichik, D., Winfree, E., Bruck, J.: Programmability of chemical reaction networks. In: Condon, A., Harel, D., Kok, J.N., Salomaa, A., Winfree, E. (eds.) *Algorithmic Bioprocesses*, pp. 543–584. Springer Berlin Heidelberg, Berlin, Heidelberg (2009)
12. Courbet, A., Amar, P., Fages, F., Renard, E., Molina, F.: Computer-aided biochemical programming of synthetic microreactors as diagnostic devices. *Molecular Systems Biology* 14(4) (2018)
13. Cousot, P., Cousot, R.: Abstract interpretation: A unified lattice model for static analysis of programs by construction or approximation of fixpoints. In: *POPL'77: Proceedings of the 6th ACM Symposium on Principles of Programming Languages*. pp. 238–252. ACM Press, New York (1977), Los Angeles
14. Craciun, G., Feinberg, M.: Multiple equilibria in complex chemical reaction networks: II. the species-reaction graph. *SIAM Journal on Applied Mathematics* 66(4), 1321–1338 (2006)
15. Fages, F., Le Guludec, G., Bournez, O., Pouly, A.: Strong Turing Completeness of Continuous Chemical Reaction Networks and Compilation of Mixed Analog-Digital Programs. In: *CMSB'17: Proceedings of the fifteen international conference on Computational Methods in Systems Biology*. Lecture Notes in Computer Science, vol. 10545, pp. 108–127. Springer-Verlag (Sep 2017)
16. Fages, F., Gay, S., Soliman, S.: Inferring reaction systems from ordinary differential equations. *Theoretical Computer Science* 599, 64–78 (Sep 2015)
17. Fages, F., Soliman, S.: Abstract interpretation and types for systems biology. *Theoretical Computer Science* 403(1), 52–70 (2008)

18. Fages, F., Soliman, S.: From reaction models to influence graphs and back: a theorem. In: Proceedings of Formal Methods in Systems Biology FMSB'08. No. 5054 in Lecture Notes in Computer Science, Springer-Verlag (Feb 2008)
19. Feinberg, M.: Mathematical aspects of mass action kinetics. In: Lapidus, L., Amundson, N.R. (eds.) Chemical Reactor Theory: A Review, chap. 1, pp. 1–78. Prentice-Hall (1977)
20. Gilbert, D., Heiner, M.: From petri nets to differential equations - an integrative approach for biochemical network analysis. In: Proceedings of ICATPN 2006. pp. 181–200. No. 4024 in Lecture Notes in Computer Science, Springer-Verlag (2006)
21. Heiner, M., Gilbert, D., Donaldson, R.: Petri nets for systems and synthetic biology. In: Bernardo, M., Degano, P., Zavattaro, G. (eds.) 8th Int. School on Formal Methods for the Design of Computer, Communication and Software Systems: Computational Systems Biology SFM'08. Lecture Notes in Computer Science, vol. 5016, pp. 215–264. Springer-Verlag, Bertinoro, Italy (Feb 2008)
22. Hucka, M., et al.: The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics* 19(4), 524–531 (2003)
23. Johnston, M.D., Anderson, D.F., Craciun, G., Brijder, R.: Conditions for extinction events in chemical reaction networks with discrete state spaces. *Journal of Mathematical Biology* 76(6), 1535–1558 (jan 2018)
24. von Kamp, A., Schuster, S.: Metatool 5.0: fast and flexible elementary modes analysis. *Bioinformatics* 22(15), 1930–1931 (2006)
25. Lebeda, F.J., Adler, M., Erickson, K., Chushak, Y.: Onset dynamics of type A botulinum neurotoxin-induced paralysis. *Journal of Pharmacokinetics and Pharmacodynamics* 35(3), 251–267 (jun 2008)
26. Nabli, F., Martinez, T., Fages, F., Soliman, S.: On enumerating minimal siphons in petri nets using CLP and SAT solvers: Theoretical and practical complexity. *Constraints* 21(2), 251–276 (2016)
27. Ovchinnikov, S.: Max-min representation of piecewise linear functions. *Contributions to Algebra and Geometry* 43(1), 297–302 (2002)
28. Peterson, J.L.: *Petri Net Theory and the Modeling of Systems*. Prentice Hall, New Jersey (1981)
29. Qian, L., Soloveichik, D., Winfree, E.: Efficient turing-universal computation with DNA polymers. In: Proc. DNA Computing and Molecular Programming. LNCS, vol. 6518, pp. 123–140. Springer-Verlag (2011)
30. Reddy, V.N., Mavrovouniotis, M.L., Liebman, M.N.: Petri net representations in metabolic pathways. In: Hunter, L., Searls, D.B., Shavlik, J.W. (eds.) Proceedings of the 1st International Conference on Intelligent Systems for Molecular Biology (ISMB). pp. 328–336. AAAI Press (1993)
31. Sackmann, A., Heiner, M., Koch, I.: Application of petri net based analysis techniques to signal transduction pathways. *BMC Bioinformatics* 7(482) (Nov 2006)
32. Schneider, F.S., Amar, P., Bahri, A., Espeut, J., Baptiste, J., Alali, M., Fages, F., Molina, F.: Biomachines for medical diagnosis. *Advanced Materials Letters* 11(4), 1535–1558 (mar 2020)
33. Senum, P., Riedel, M.: Rate-independent constructs for chemical computation. *PLOS One* 6(6), e21414 (2011)
34. Soliman, S.: Invariants and other structural properties of biochemical models as a constraint satisfaction problem. *Algorithms for Molecular Biology* 7(15) (May 2012)