Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Input-Output instability patterns of Chemical Reaction Networks

David Angeli
Dept. of Electrical and Electronic Engineering, Imperial College, London
Dip. di Sistemi e Informatica, Università di Firenze

Abstract— **This paper describes a criterion for qualitative analysis of open Chemical Reaction Networks endowed with mass-action kinetics. The method can be applied to an extremely broad class of such open networks, and returns, as an outcome, a classification of the possible dynamical behaviors that are compatible with the network structure, by classifying each species as bounded, converging to 0 or diverging to ∞.**

I. Introduction and Motivations

This paper deals with questions of external stability for systems with inputs. Sometimes the species concentrations of a reaction network will remain bounded provided bounded inputs are applied. This property is usually referred to as Bounded Input Bounded State Stability (BIBS for short). On the other hand, in case of systems which are not BIBS stable, we would like to characterize which components of the state may potentially diverge.

While precise definitions are provided later on, it is worth mentioning that standard assumptions often result in compartmental dynamical models, at least in the absence of external inflows. The qualitative analysis of such systems typically poses challenging questions. Indeed if we consider the action of external inflows the resulting dynamical behavior crucially depends upon a balance between inflows and outflows. Unlike inflow rates, which are typically a priori known (and postulated for instance to be constant), output rates are, more often, a function of the state and their influence can hardly be quantified (at least a priori), thus making qualitative analysis a rather challenging task. Indeed, such questions arise across several domains of applications and apparently distant modeling frameworks: not only positive compartmental systems (see for instance [8]), but also Petri Nets (in the context of Discrete Event Systems [4]). For instance, in the case of logic Petri Nets, consideration of inflows (and outflows) typically leads to infinite reachable spaces, showing that the notion of boundedness therein adopted is somewhat inappropriate for studying such stability issues. On the other hand, consideration of Stochastic Timed Petri Nets (with exponential clocks), induces countable Markov Chains, and therefore, claiming existence of a steady-state probability distribution for constant inflows might be difficult in general, let alone proving boundedness of momenta of such distributions. In the context of compartmental dynamical systems, it is nonlinearity of the corresponding governing equations that poses the major challenges. Indeed, for linear compartmental systems, BIBS stability is well understood and, in the irreducible case, equivalent to having an Hurwitz matrix A defining the systems equations. Moreover, all state variables are either simultaneously bounded or simultaneously unbounded, due to existence of a strictly positive dominant eigenvector of A (the so called Perron-Frobenius eigenvector). The nonlinear case, on the contrary, appears to be much more delicate and only weak sufficient conditions are typically available for claiming BIBS stability of general purpose systems (for instance postulating existence of outflows in every single state compartment whose rates, in addition, tend to infinity as the corresponding state variables grow unbounded, see [3]). Indeed, the most interesting cases arise when only part of the variables are directly connected to an outflow of the system.

This paper describes the working principles of an algorithm for qualitative analysis of open Chemical Reaction Networks endowed with mass-action kinetics. Space constraints do not allow a description of the algorithm itself, which is based on branch and bound techniques, and will be detailed in a forthcoming journal publication. Notice that, consistently with our previous remarks, the type of kinetics considered might have a crucial impact on the overall behavior. Indeed, as confirmed by simulations, even the specific values of parameters considered may heavily affect the outcome of experiments. On the positive side, the algorithm can be applied to an extremely broad class of such open networks, and returns, as an outcome, a classification of the possible dynamical behaviors that are compatible with the network topology. Remarkably, simulations on non-trivial examples showed that, as the kinetic constants of the network are varied, all the predicted behaviors could be observed.

II. Background on Chemical Reaction Networks

In this section we closely follow the definitions given in [2]. In mathematical terms, a *chemical reaction network* ("CRN", for short) is a list of chemical reactions $\mathcal{R}_i$, where the index i takes values in $\mathcal{R} := \{1, 2, \ldots, n_r\}$. In particular, using the formalism from chemical networks theory (see [5] for an interesting introduction to the topic), we consider a set of chemical species S_j, $j \in \{1, 2, \ldots n_s\} := \mathcal{S}$ which are the compounds taking part in the reactions. Individual chemical reactions are then denoted as follows:

$$\mathcal{R}_i : \quad \sum_{j\in\mathcal{S}} \alpha_{ij} S_j \to \sum_{j\in\mathcal{S}} \beta_{ij} S_j \tag{1}$$

where the α_{ij} and β_{ij} are nonnegative integers called the *stoichiometry coefficients*. The compounds on the left-hand side are usually referred to as the *reactants*, and the ones on the right-hand side are called the *products*, of the reaction. Informally speaking, the forward arrow means that the transformation of reactants into products only happens in the direction of the arrow. It is worth pointing out that we allow chemical reactions in which the right-hand and left-hand side are actually empty. This case corresponds, from a physical point of view, to *outflows* and *inflows* of the chemical reaction. The modeling of inflows, viz. of species whose concentration is a priori specified and unaffected by the reaction network considered, is achieved by allowing *potentially time-varying* reaction rates.

As usual, we arrange the stoichiometry coefficients into an $n_s \times n_r$ matrix, called the *stoichiometry matrix* Γ, defined as follows:

$$[\Gamma]_{ji} = \beta_{ij} - \alpha_{ij}, \tag{2}$$

for all $i \in \mathcal{R}$ and all $j \in \mathcal{S}$ (notice the reversal of indexes). This will be later used in order to synthetically write the differential equation associated to a given chemical network. Notice that we allow Γ to have columns which differ only by their sign; this happens when there exist reversible reactions in the network.

We discuss, next, how the speed of reactions is affected by the concentrations of the different species. Each chemical reaction takes place continuously in time, at its own rate, possibly time-varying, which is assumed to be only a function of the concentration of the species taking part in it. In order to make this more precise, we define the vector $S = [S_1, S_2, \dots S_{n_s}]'$ of species concentrations and, as a function of it, the vector of reaction rates

$$R(S,t) := [R_1(S,t), R_2(S,t), \dots R_{n_r}(S,t)]' .$$

Although more general versions can be envisioned, the one discussed throughout this paper, only deals with a special form of reaction rates. The so called *mass-action kinetics*, which correspond to the following expression:

$$R_i(S,t) = k_i(t) \prod_{j=1}^{n_s} S_j^{\alpha_{ij}} \text{ for all } i = 1, \dots, n_r.$$

With the above notations, the chemical reaction network can synthetically be described by the following system of differential equations:

$$\dot{S}(t) = \Gamma\, R(S(t), t). \tag{3}$$

III. A Petri Net's interpretation

We associate to a CRN a bipartite directed graph (i.e. a directed graph with two types of nodes) with weighted edges, called the *species-reaction Petri net*, or SR-net for short. Mathematically, this is a quadruple

$$(V_S, V_R, E, W) ,$$

where V_S is a finite set of nodes, each one associated to a species, V_R similarly is a finite set of nodes (disjoint from V_S) corresponding to reactions, and E is a set of edges as described below. (We often write S or V_S interchangeably, or R instead of V_R, by identifying species or reactions with their respective indexes; the context should make the meaning clear.) The set of all nodes is also denoted by $V \doteq V_R \cup V_S$.

The edge set $E \subset V \times V$ is defined as follows. Whenever a certain reaction R_i belongs to the CRN:

$$\sum_{j \in \mathcal{S}} \alpha_{ij} S_j \quad \to \quad \sum_{j \in \mathcal{S}} \beta_{ij} S_j \, , \tag{4}$$

we draw an edge from $S_j \in V_S$ to $R_i \in V_R$ for all S_j's such that $\alpha_{ij} > 0$. That is, $(S_j, R_i) \in E$ iff $\alpha_{ij} > 0$, and we say in this case that R_i is an *output reaction for* S_j. Similarly, we draw an edge from $R_i \in V_R$ to every $S_j \in V_S$ such that $\beta_{ij} > 0$. That is, $(R_i, S_j) \in E$ whenever $\beta_{ij} > 0$, and we say in this case that R_i is an *input reaction for* S_j.

Notice that edges only connect species to reactions and vice versa, but never connect two species or two reactions.

The last element to fully define the Petri net is the function $W : E \to \mathbb{N}$, which associates to each edge a positive integer according to the rule:

$$W(S_j, R_i) = \alpha_{ij} \quad \text{and} \quad W(R_i, S_j) = \beta_{ij} \, .$$

Therefore, the so called stoichiometry matrix is nothing more than the classical *incidence matrix* of the associated Petri Net. To illustrate these concepts, consider the following enzymatic reaction network:

$$\begin{array}{c} S_0 + E \leftrightarrow ES_0 \to S_1 + E \leftrightarrow ES_1 \to S_2 + E \\ S_2 + F \leftrightarrow FS_2 \to S_1 + F \leftrightarrow FS_1 \to S_0 + E \\ \emptyset \to S_0 \qquad S_2 \to \emptyset \end{array} \tag{5}$$

This is a short hand notation for the list of chemical reactions reported in the Table below:

$\mathcal{R}_1$	$S_0 + E \to ES_0$	$\mathcal{R}_2$	$ES_0 \to S_0 + E$
$\mathcal{R}_3$	$ES_0 \to S_1 + E$	$\mathcal{R}_4$	$S_1 + E \to ES_1$
$\mathcal{R}_5$	$ES_1 \to S_1 + E$	$\mathcal{R}_6$	$ES_1 \to S_2 + E$
$\mathcal{R}_7$	$S_2 + F \to FS_2$	$\mathcal{R}_8$	$FS_2 \to S_2 + F$
$\mathcal{R}_9$	$FS_2 \to S_1 + F$	$\mathcal{R}_{10}$	$S_1 + F \to FS_1$
$\mathcal{R}_{11}$	$FS_1 \to S_1 + F$	$\mathcal{R}_{12}$	$FS_1 \to S_0 + F$
$\mathcal{R}_{13}$	$\emptyset \to S_0$	$\mathcal{R}_{14}$	$S_2 \to \emptyset$

This network can be graphically represented as in Fig. 1, with the standard convention that reactions (transitions in Petri Net terminology) are associated to bars (or boxes), while species (places) are represented by circular nodes.

Several other definitions which are commonly used in the Petri net literature will be of interest in the following. We say that a row or column vector v is non-negative, and we denote it by $v \succeq 0$ if it is so entry-wise. We write $v \succ 0$ if $v \succeq 0$ and $v \neq 0$. A stronger notion is instead $v \gg 0$, which indicates $v_i > 0$ for all i.

Definition 3.1: A *P-invariant* (respectively T-invariant) is any integer row (column) vector $c \succ 0$ such that $c\Gamma = 0$ (respectively $\Gamma\, c = 0$). Its *support* is the set of indices $\{i \in V_S : c_i > 0\}$. □

Notice that each P-invariant for the system (3) corresponds to some non-negative linear first integral, that is, to a linear

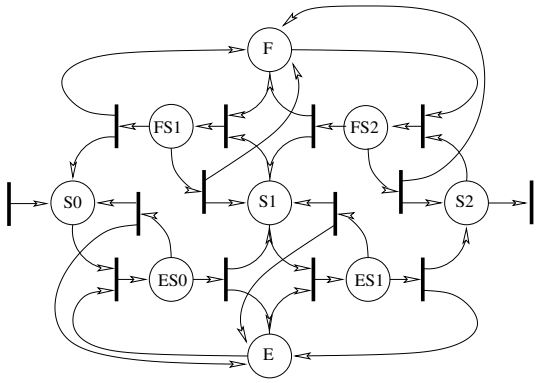

Fig. 1. Petri Net associated to a chemical reaction network

function $S \mapsto cS$ such that $(d/dt)cS(t) \equiv 0$ along all solutions of (3) A *P-decreasing* (P-increasing) vector is any integer row vector $c \succ 0$ such that $c\,\Gamma \preceq 0$, ($c\,\Gamma \succeq 0$).

Our goal is to analyze the system irrespectively of the values of unknown parameters and of rates at which chemicals are externally provided to the system. This could in general be a rather formidable task, and for this reason we adopt the following assumption.

Key Hypothesis: We assume that the solutions of (3) are such that each species $S_i(t)$, $i \in \mathcal{S}$ (and similarly each reaction rate $R_j(S(t),t)$), fulfills the following trichotomy:

1) $\lim_{t\to+\infty} S_i(t) = 0$;
2) $$0 < \liminf_{t\to+\infty} S_i(t) \leq \limsup_{t\to+\infty} S_i(t) < +\infty$$
3) $\lim_{t\to+\infty} S_i(t) = +\infty$.

Of course this assumption does not rule out the possibility of having complex behaviors (such as chaotic or oscillating concentrations), and, at the same time, allows for chemical species to also vanish to 0 or grow to ∞. What it is ruled out, instead, is the possibility of exotic solutions which along some time subsequence converge to 0, while along others have a finite or infinite limit. Indeed, while it is clear that dynamical systems can exhibit such behaviors, simulations with meaningful practical examples, like those arising in biochemistry, show that such assumption is indeed rather mild and solutions satisfying the Key Hypothesis are, by far, the most commonly met. Hence, it appears worthwhile to pursue an analysis of the possible network's behaviors under the Key Hypothesis, which, on the contrary, from the analytic point of view constitutes a major reduction in complexity.

When multiple simultaneous reactions are taken into account, the analysis may quickly become very cumbersome. Indeed, if the network is covered by P-vectors, this entails monotonicity of some integer combinations of state variables and/or existence of conservation laws, which can be used to infer the asymptotic behavior of the system. Furthermore, such conditions are structural and do not depend upon systems parameters. In most cases, however, there are not enough a priori guaranteed monotonicity conditions, and the final qualitative behavior critically depends upon the balance between inflows and outflows. Under such premises, the Key Hypothesis, may provide useful hints to single out the possible qualitative dynamical behaviors associated to each chemical species. Our main idea is to label each species with a 0, 1 or ∞, according to whether item 1. ,2. or 3. of the Key Hypothesis holds. Hopefully, of all possible ways of carrying out such labelings, (namely $3^9 = 19683$ for the network under consideration), only few of them will be compatible with the topology and the structural invariants of the network. In the following we would like to illustrate ideas to design an algorithm for the explicit enumeration of all such meaningful labelings. To this end, the following Lemma classifying the possible dynamical behaviors of scalar linear time-varying systems, is crucial:

Lemma 3.2: Consider the time-varying scalar linear system given below:

$$\dot{x}(t) = -a(t)x(t) + b(t) \tag{6}$$

where $a(t)$ and $b(t)$ are continuous non-negative scalar functions of time, satisfying the trichotomy expressed in the Key Hypothesis. Then, if $x(t)$ also fulfills the Key Hypothesis, the only compatible dynamical behaviors are listed in the table below:

$a(t)\backslash b(t)$	0	1	∞
0	u	∞	∞
1	0	1	∞
∞	0	0	u

where u stands for *undetermined.*
The proof of Lemma 3.2 relies on the explicit solution formula available for scalar linear time-varying equations (6) and is omitted for space reasons.

Thanks to mass-action kinetics and *in the absence of autocatalytic reactions*, viz. reactions comprising an identical species among reactants and products, it is easy to recognize the structure of equation (6) in the update equation of each species. In particular, the coefficient $b(t)$ can be simply computed as the sum of the rates of incoming reactions, while $a(t)$ is the sum of the rates of outgoing reactions divided by the concentrations of the species itself (this can be always done without harm, since a reactant species is also a factor of the corresponding reaction rate). Let us carry out this simple computation just once, for illustrating purposes, relatively to the species $S_0(t)$ of network (5). Its update equation reads:

$$\dot{S}_0(t) = -k_1 S_0(t) \cdot E(t) + k_2 E S_0(t) + k_{12} F S_1(t) + k_{13}$$

Hence, the coefficients $a(t)$ and $b(t)$ are given respectively by:

$$a(t) = k_1 E(t) \quad b(t) = k_2 E S_0(t) + k_{12} F S_1(t) + k_{13}.$$

Thus, labels associated to $a(t)$ and $b(t)$ can be computed as a function of the labels attached to the species. To this end, it is enough to define a suitable set of rules, defining how the product and sum operations behave as a function

of individual labels. It is straightforward to verify that the tables given below can be employed to compute the labels attached to the sum and product of functions:

$+$	0	1	∞
0	0	1	∞
1	1	1	∞
∞	∞	∞	∞

$\cdot$	0	1	∞
0	0	0	u
1	0	1	∞
∞	u	∞	∞

where u stands for *undetermined*.

The idea is to test each labeling of all species by computing potential labels associated to the corresponding reaction rates and finally those associated to the coefficients $a(t)$ and $b(t)$ in the update equation of individual species. Then, thanks to Lemma 3.2, we can verify consistence of the label of species S_i with that obtained from the table of the basis of the resulting $a(t)$ and $b(t)$ coefficients.

Another significant consistency check can be performed on the basis of P-invariant P-decreasing and P-increasing vectors associated to the Net.

To this end for a given vector v, we denote by $\mathcal{O}(v)$ the set of out-flowing reactions, that is: $\mathcal{O}(v) = \{R_j : v'\Gamma e_j < 0\}$ (with e_j the j-th element of the canonical basis of $\mathbb{R}^{n_r}$) and by $\mathcal{I}(v)$ the set of in-flowing reactions, that is: $\mathcal{I}(v) = \{R_j : v'\Gamma e_j > 0\}$. With this notation we are able to clarify what kind of consistency checks can be performed on the basis of P-invariants.

Lemma 3.3: Let $\Sigma \subset \mathcal{S}$ be the support of a P-vector v. Let $S(t)$ be a solution of (3) satisfying the Key Hypothesis and $L(S_i)$, $L(R_j)$ be respectively the labels associated to species S_i and reaction R_j. Then, label patterns fulfill:

- P-invariant: $\forall\, S_i \in \Sigma, L(S_i) \in \{0, 1\}$
- P-decreasing vector: $\forall\, S_i \in \Sigma, L(S_i) \in \{0, 1\}$ and $\forall\, R_j \in \mathcal{O}(v), L(R_j) = 0$
- P-increasing vector: $\exists\, S_i \in \Sigma : L(S_i) \in \{1, \infty\}$ and $\big(\nexists\, S_i \in \Sigma : L(S_i) = \infty\big) \Rightarrow \big(\forall\, R_j \in \mathcal{I}(v),\ L(R_j) = 0\big)$

IV. Asymptotic P-vectors and T-vectors analysis

As a by-product of the species labeling procedure, one also obtains a vector of corresponding labels attached to reactions. In this respect, it is not unfrequent in the search of labelings which are consistent with the network topology, to find reactions with a 0 label. This implies that, at least asymptotically, that reaction can be regarded as switched off or not being part of the network any longer. Therefore, for each given species and reaction labeling, it is meaningful to define the set $\mathcal{R}_+ \subset \mathcal{R}$ of reactions whose attached label is different from 0. Accordingly, we may reduce the dimension of the stoichiometry matrix Γ by only keeping the columns corresponding to reactions in $\mathcal{R}^+$. We call this new matrix $\tilde{\Gamma}$. This may alter in a significant way the P-vectors associated to the net, so that, indeed, new P-decreasing and P-invariant vectors can possibly be found. The asymptotic P-vectors analysis is based on this very idea: namely, once a candidate labeling is provided, compute the P-decreasing vectors associated with the reduced network $\tilde{\Gamma}$. Then, make sure that the following constraints are fulfilled:

1) each P-decreasing vector of $\tilde{\Gamma}$ is also a P-invariant
2) for each P-increasing vector of $\tilde{\Gamma}$ which is not also a P-invariant, make sure that there exists at least one species in its support labeled with ∞.
3) the reduced network is covered by T-increasing vectors.

We show next, by contradiction, that P-decreasing vectors of $\tilde{\Gamma}$ which are not also P-invariant cannot exists. Let $c \succ 0$ be such that $c\tilde{\Gamma} \prec 0$. Let us consider the derivative of $cS(t)$ along solutions of (3); by definition of $\tilde{\Gamma}$ we have indeed:

$$c\dot{S}(t) = c\left[\sum_{j:\mathcal{R}_j=0} (\beta_j - \alpha_j) k_j S^{\alpha_j}(t)\right] + c\tilde{\Gamma}\tilde{R}(S(t), t) \quad (7)$$

where $\tilde{R}(S, t)$ corresponds to the vector of reaction rates obtained from $R(S, t)$ by removing reactions with 0 asymptotic rate. Indeed, the first term at the right hand side of (7) tends to 0 asymptotically, so that, for some $T > 0$ there exists $k > 0$ with the property that $c\dot{S}(t) \leq -k$ for all $t \geq T$. This yields a contradiction, since it implies $cS(t) \leq cS(T) - k(t - T)$, and hence $cS(t) \to -\infty$ as $t \to +\infty$ which is absurd. A symmetric argument shows that, for each P-increasing vector which is not a P-semiflow we obtain $c\dot{S}(t) \geq k$ for all $t \geq T$ for positive k and T. Hence $cS(t) \to +\infty$ as $t \to +\infty$, thus implying that at least one place of the support of c should have an ∞ label.

Dually, we may consider the asymptotic T-vectors associated to the reduced network. Indeed, partitioning Γ as $[\tilde{\Gamma}, \Gamma_0]$ and $R(S, t) = [\tilde{R}(S, t)', R_0(S, t)']'$, with Γ_0 contains the columns of Γ associated to reactions with 0 asymptotic rate, and $R_0(S, t)$ the expression of the corresponding kinetics, we have, by simple integration: $S(t) - S(0) = \tilde{\Gamma} \int_0^t \tilde{R}(S(\tau), \tau)\, d\tau + \Gamma_0 \int_0^t R_0(S(\tau), \tau)\, d\tau$. Hence, dividing by t and letting $t \to +\infty$ in both sides, possibly along a subsequence, for the sake of limit existence (either finite or infinite), we obtain $\bar{S} = \tilde{\Gamma}\bar{R}$ where $\bar{S} = \lim_{n\to+\infty}[S(t_n) - S(0)]/t_n \succeq 0$ and $\bar{R}$ belongs to the set of asymptotic averages of $\tilde{R}(S(t), t)$, and hence, by construction it fulfills $\bar{R} \gg 0$. Indeed, $\tilde{\Gamma}\bar{R} \succeq 0$ which shows that the reduced network is covered by T-increasing vectors.

V. Simulation results

In the present Section we illustrate the outcome of the analyzer for the special enzymatic reaction network considered in (5). To express our results we list the species in the following order: $[S_0, S_1, S_2, E, F, ES_0, ES_1, FS_2, FS_1]$. We run the analyzer by letting $\mathcal{R}_{13}$ have a 1 label, (which stands for a persistently active input, for instance a constant one). The analysis shows that only 6 labelings are compatible with the topology considered here, namely:

Species	I	II	III	IV	V	VI
S_0	∞	∞	1	∞	∞	∞
S_1	1	1	1	∞	∞	∞
S_2	0	1	1	0	1	1
E	0	0	1	0	0	0
F	1	1	1	0	0	0
ES_0	1	1	1	1	0	1
ES_1	0	0	1	0	1	1
FS_2	0	1	1	0	0	0
FS_1	1	1	1	1	1	1

Simulations show that the most intuitive situation, scenario III, in which all species stay bounded and bounded away from zero, only appears for sufficiently small input values. As the input grows different type of phenomena occur. In particular, an excess of S_0, leads to a rapid combination with E molecules, which, as a result, tend to disappear so that, overall, the conversion of S_0 to S_1 and ultimately S_2 may slow down and asymptotically stop. Let us concentrate next on scenario V. The labels corresponding to species concentrations and reaction rates are displayed in Fig. 3. Notice that arcs have been drawn with lines of different thickness according to the label of the corresponding reaction, namely thin line for reactions with a 0 label and thick line for reactions with a 1 label. In this scenario, reaction $\mathcal{R}_{14}$, which is an outflow of the system, is asymptotically active. Hence, a positive average outflow spills out of the chemical reaction network, which, intuitively, entails a positive average inflow and a path which connects such inflow (namely $\mathcal{R}_{13}$), to the transition $\mathcal{R}_{14}$. It is easy to verify by inspection that such a path does not exist in scenario V, as $\mathcal{R}_{14}$ is not reachable from $\mathcal{R}_{13}$ by only following thick edges. Indeed, the reduced net obtained by taking out transitions with 0 label is not covered by T-increasing vectors. Then, this scenario can be discarded. Alternatively, the scenario could be discarded by performing an asymptotic P-invariant analysis. Namely, by verifying that $[0,1,0,0,0,0,1,0,1,0]$ is a P-decreasing vector for the matrix $\tilde{\Gamma}$ but it is not at the same time a P-invariant. Hence, it violates both conditions exposed in Section IV.

We consider next scenario II. The corresponding labels are displayed in Fig. 2. Notice that, in this hypothetical scenario, reaction $\mathcal{R}_{14}$ is active, and indeed a path with positive reaction average rates connects $\mathcal{R}_{13}$ to the outflow $\mathcal{R}_{14}$. However, a formal verification shows that the reduced network is not covered by T-increasing vectors. Hence, this scenario can also be discarded. Indeed, from the physical point of view, S_2 represents a doubly phosphorilated version of S_0, and hence, only species containing S_0 can be accounted for such transfer flow. Enzyme F, on the other hand, favors certain chemical reactions but does not carry itself molecules of S, whereas the only active path from $\mathcal{R}_{13}$ to $\mathcal{R}_{14}$ goes precisely through F. Hence, we are only left with 4 possible distinct scenarios. Simulations show that all of them are actually observed for suitable choices of kinetic constants and initial conditions. The most common situation, at least for a random sampling of kinetic parameters with uniform distributions in a logarithmic scale, are scenarios I and VI. Scenario IV is rather unfrequent in the considered parameters space, nevertheless it appeared in a small fraction of our simulations.

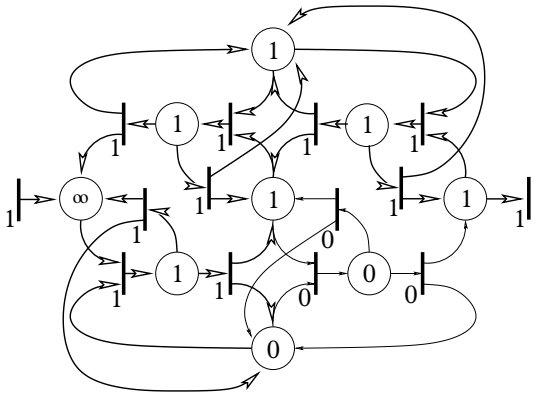

Fig. 2. Scenario II: species and reaction labels

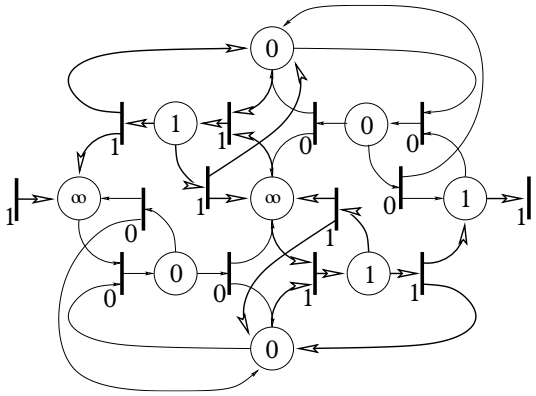

Fig. 3. Scenario V: species and reaction labels

VI. Discussion and Conclusions

A. The analysis technique

We have presented an analysis technique for *open* chemical reaction networks which enables, on the basis of purely topological and structural hypothesis, to rule out dynamical behaviors which are not compatible with the given structure. The analysis technique assumes some a priori knowledge on the possible systems trajectories (the so called Key Hypothesis) which states three alternatives for the asymptotic properties of individual concentrations of each chemical species. Though not easily justifiable from a purely mathematical point of view, such assumptions appear to be rather mild in practical examples arising from biochemistry. Remarkably, while each concentration is free to exhibit any out of 3 prespecified asymptotic features, the overall network turns out to be highly constrained, by its own topology and regardless of potentially unknown parameters entering the equations, in terms of the possible combinations of such individual behaviors.

B. Potential impact

The analysis proposed may have potential benefits in many relevant challenges that experimental biologists are

constantly faced with; we list some of the most obvious ones, without aim of completeness:

1) Validation or falsification of modeling hypotheses: a very difficult task and a preliminary step in trying to estimate quantitative mathematical models of biological networks appears to be the formulation of a plausible network layout which could justify the experimental data. Our tool appears to give theoretically sound criteria for ruling out topologies which are not compatible with experimental data observed. This can be done even before attempting to estimate the actual parameters characterizing the system, which is usually a very time-consuming and costly task and even on the basis of very partial data, such as measures of only a subset of the species which are believed to interact within a given chemical network.
2) Iterative formulation of hypothesis: though a reverse algorithm has not been developed yet, namely an algorithm which starting from a given observed experimental labels vector (or why now a set of them) goes about extracting all the possible networks topologies which are consistent with it, it is clear that our consistency check should provide an insight in the process of iteratively upgrading hypothetical networks topologies to match the data observed.
3) Experiment design: what are the meaningful experiments which will allow to discriminate between two equally plausible hypothesis. Not all experiments have this potential; however, if among the consistent behaviors there exists at least one which is respectively allowed by one topology and not by the other it might be a good idea to try to actually experimentally observe this behavior on the real system. If it is indeed observed, then one hypothesis can be discarded.
4) Synthesis of chemical networks with prescribed behavior: as usual a powerful analysis can be the starting point for a synthesis process. Of course the major difficulty will, in this respect, be the possibility of actually tuning the kinetic parameters in order to reproduce a desired behavior; a preliminary step, in any case, would be the individuation of the required network topology.

C. Relation to previous literature on closed CRN

When applied to closed CRN, the algorithm only returns 0 and 1 labels as an outcome (indeed boundedness of solutions is a priori guaranteed). Since unboundedness is not an issue for closed systems, processing a CRN with this algorithm amounts to studying persistence of a chemical reaction network, viz. non-extinction (not even asymptotically) of all its chemical species. This property, in the continuous set-up we are dealing with, though formulated in terms of species concentrations, rather than reaction rates, can be thought of as an analogue of Liveness for Petri Nets. See [1] for considerations along these lines. Indeed, it is possible to show that if a set of 0 labels exist as an output of the algorithm, then this set is necessarily a siphon of the associated Petri Net. Indeed, this result is consistent with the main result in [2], where it is shown that whenever solutions approach the boundary of the positive orthant (even without assuming the Key Hypothesis), then the corresponding set of variables which tend to 0 needs to be a siphon for the associated Petri Net. In particular, it must be a *critical siphon*, that is a siphon, which does not include the support of any P-semiflows of the net. Accordingly, a siphon which includes the support of a P-semiflow is called a *structurally non-emptiable siphon* and none of its species can ever asymptotically decay to zero (notice the similarity with the so called siphon-trap condition). Hence, sufficient conditions for persistence clearly show that, at least in the case of closed systems, remotion of the Key Hypothesis does not include other types of convergence patterns to zero as far as state variables are concerned.

D. Description of the Algorithm

Efficient implementation of the algorithm is based on the selective exploration of the tree of all possible labelings. In particular, rather than proceeding with the computation of all possible labelings and some a posteriori check of consistency, we gradually update the currently explored configuration by making use of indeterminate labels, viz. labels corresponding to the following 7 different possibilities: 0, 1, ∞, 01, 1∞, 0∞, 01∞. Those involving more than one symbol are meant to be associated to state variables for which more than one possible asymptotic behavior is still plausible and has to be simultaneously taken into account. The tree is explored thanks to a branch and bound technique which performs consistency checks all the way down from the root to the leaves. In such a way, incompatible labelings are discarded much earlier in the construction of the tree.

REFERENCES

[1] D Angeli, P de Leenheer, E.D. Sontag, "A Petri net approach to persistence analysis in chemical reaction networks" *Biology and Control Theory: Current Challenges (Lecture Notes in Control and Information Sciences Volume 357)*, Springer-Verlag, Berlin (2007), pp. 181-216.

[2] D. Angeli, P. de Leenheer, E.D. Sontag, "A Petri net approach to the study of persistence in chemical reaction networks"Mathematical Biosciences (2007),

[3] G. Bastin, "Issues in modelling and control of mass balance systems", in *Stability and Stabilization of Nonlinear Systems*, Edited by D. Aeyels, F. Lamnabhi-Laguarrigue and A.J. van der Schaft, *Lecture Notes in Control and Information Sciences*, Springer, Vol. 246, pp. 53-74, 1999.

[4] C.G. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*, Springer, 2008

[5] M. Feinberg, "Lectures on chemical reaction networks" *Lectures at the Mathematics Research Center, University of Wisconsin*, 1979.

[6] R. Hofestädt, "A Petri net application to model metabolic processes" *Syst. Anal. Mod. Simul.* **16** (1994), pp. 113-122.

[7] F.J.M. Horn, "The dynamics of open reaction systems," in *Mathematical aspects of chemical and biochemical problems and quantum chemistry (Proc. SIAM-AMS Sympos. Appl. Math., New York, 1974)*, pp. 125-137. SIAM-AMS Proceedings, Vol. VIII, Amer. Math. Soc., Providence, 1974.

[8] J.A. Jacquez and C.P. Simon, "Qualitative theory of compartmental systems", *SIAM Review*, Vol. 35, N.1, pp. 43-79, 1993.