



Københavns Universitet



## Intermediates and Generic Convergence to Equilibria

Marcondes de Freitas, Michael; Wiuf, Carsten; Feliu, Elisenda

*Published in:*  
Bulletin of Mathematical Biology

*DOI:*  
[10.1007/s11538-017-0303-4](https://doi.org/10.1007/s11538-017-0303-4)

*Publication date:*  
2017

*Document Version*  
Early version, also known as pre-print

*Citation for published version (APA):*  
Marcondes de Freitas, M., Wiuf, C., & Feliu, E. (2017). Intermediates and Generic Convergence to Equilibria. *Bulletin of Mathematical Biology*, 79(7), 1662–1686. <https://doi.org/10.1007/s11538-017-0303-4>

# Intermediates and Generic Convergence to Equilibria

Michael Marcondes de Freitas, Carsten Wiuf, and Elisenda Feliu

Department of Mathematical Sciences, University of Copenhagen

July 1, 2016

## Abstract

Known graphical conditions for the generic or global convergence to equilibria of the dynamical system arising from a reaction network are shown to be invariant under the so-called successive removal of intermediates, a systematic procedure to simplify the network, making the graphical conditions easier to check.

**Keywords.** Reaction Network Theory · Model Reduction · SR-graph · Monotonicity in Reaction Coordinates

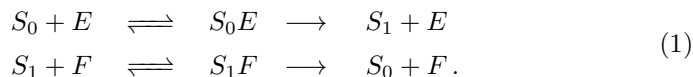
## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Reaction Networks</b>	<b>4</b>
2.1	Basic Formalism . . . . .	4
2.2	Graph Conditions for Generic Convergence . . . . .	6
<b>3</b>	<b>Main Results</b>	<b>9</b>
3.1	Removal of Intermediates . . . . .	10
3.2	Invariance under the Removal of an Intermediate . . . . .	12
3.3	Examples . . . . .	12
3.4	Further Comments on Definition 2 . . . . .	16
<b>4</b>	<b>Proofs of Theorems 1 and 2</b>	<b>17</b>
4.1	Proof of Theorem 1( <i>i</i> ) . . . . .	17
4.2	Proof of Theorem 1( <i>ii</i> ) . . . . .	18
4.3	Proof of Theorem 2 . . . . .	19

# 1 Introduction

In recent years many works in reaction network theory have been concerned with the idea of model reduction or simplification. This interest is expressed along various lines of investigation. One direction is the natural problem of providing simpler models to describe or explain the same biochemical phenomenon [12]. Another dimension is the consolidation of known model simplification techniques typically justified and applied *ad hoc*, such as quasi-steady state approximations [9, 5], into a formal procedure [13]. A third line of inquiry contemplates whether certain qualitative properties of a reaction network, for instance, number of steady states [6] or the property of persistence [11], are invariant under a simplification procedure. This work fits within this last category. The qualitative property of interest is generic convergence to equilibria—the property that almost every solution within each stoichiometric compatibility class approach the set of equilibria—and the model simplification procedure is the successive removal of intermediates (in the sense of Definition 2).

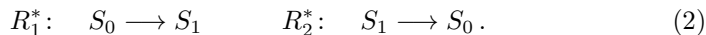
To illustrate our contribution, consider the one-site phosphorylation mechanism modeled by the reaction network



In this mechanism,  $S_0$  and  $S_1$  are, respectively, the dephosphorylated and phosphorylated forms of some protein. The phosphorylation and dephosphorylation reactions are catalyzed by a kinase  $E$  and a phosphatase  $F$ . Intermediate steps in the process during which protein and catalyst are bound to one another are captured in  $S_0E$  and  $S_1F$ . Activation/deactivation motifs such as this one appear in many important intracellular signaling processes regulating cell proliferation, differentiation and apoptosis in eukaryotes ranging from yeast to mammals [17].

In [3], sufficient graphical conditions for a reaction network to exhibit generic convergence to equilibria (within each stoichiometric compatibility class) were given. The technique consists of checking that the R-graph of the network is such that every simple loop has an even number of negative edges, a property known as the positive loop property, and that there exists a directed path between any two reaction nodes in the directed SR-graph. For the one-site phosphorylation mechanism above, the directed SR-graph and the R-graph are displayed in Figure 1, and one can readily see that they satisfy the aforementioned conditions.

By successively removing the intermediates  $S_0E$  and  $S_1F$ , what basically consists of “collapsing” the reaction paths through them, followed by canceling out the “catalysts”  $E$  and  $F$ , which appear on both sides of their respective emerging reactions with the same stoichiometric coefficient, we obtain the simplified substrate network



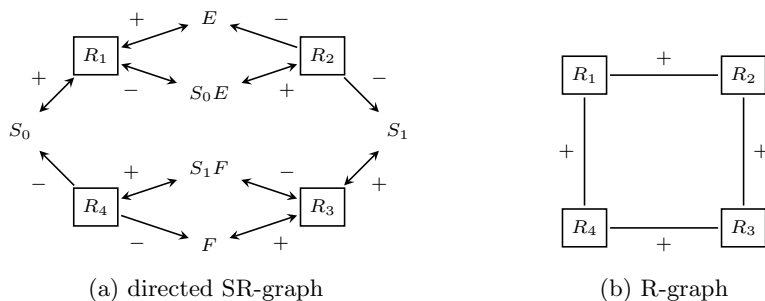


Figure 1: The directed SR- and the R-graph of (1).

For this simplified network, the directed SR-graph and the R-graph are much simpler (Figure 2), and the conditions for generic convergence are much easier to check. (The reason for not writing this as a single reversible reaction will become clearer when we introduce our working reaction network formalism in the next section.)



Figure 2: The directed SR- and the R-graph of (2).

In what follows, we will show that R-strong connectivity of the directed SR-graph and the positive loop property of the R-graph are always invariant under the successive removal of intermediates, meaning that the reduced network has them if, and only if the original one does also, as illustrated in the example above. Thus, the conditions for the original model can be checked in the often times much simpler reduced model. Therefore, although this “invariance under reduction” feature might be useful in the context of finding simpler models to describe the same observed phenomenon, it is also interesting on its own as a mathematical tool to analyze large, complicated models, even if the network obtained through the reduction procedure might not necessarily be understood to be biologically meaningful.

The approach to generic convergence to equilibria in [3] is based upon the monotone systems theory of M. W. Hirsch [7, 15, 8]. The reader familiar with that theory will likely notice the connection, although most of the details have been deliberately hidden in our presentation by framing all concepts and results pertaining to monotonicity directly in terms of the graphical conditions given in [3].

This paper is organized as follows. In Section 2 we introduce our basic notation and working definition of reaction network, then review the graphical

conditions for generic convergence to equilibria of [3]. In Section 3 we describe a systematic procedure to obtain a reduced reaction network by successively removing intermediates from a given network. We then state our main results concerning the invariance of the aforementioned conditions for generic/global convergence under this procedure, and apply them to several examples in the recent reaction network literature. The last section is devoted to the technical details of the proofs of our main results.

## 2 Reaction Networks

In what follows, we denote the set of nonnegative real numbers by  $\mathbb{R}_{\geq 0}$ , and denote the set of strictly positive real (respectively, integer) numbers by  $\mathbb{R}_{> 0}$  (respectively,  $\mathbb{Z}_{> 0}$ ). Given  $n \in \mathbb{Z}_{> 0}$ , we write  $[n] := \{1, \dots, n\}$ . By convention,  $[0] := \emptyset$ . For each  $a \in \mathbb{R}$ ,

$$\text{sign } a := \begin{cases} 1, & \text{if } a > 0 \\ 0, & \text{if } a = 0 \\ -1, & \text{if } a < 0. \end{cases}$$

### 2.1 Basic Formalism

We start by introducing our working definition of reaction network. A *complex* over a nonempty, finite set  $\mathcal{S} = \{S_1, \dots, S_n\}$  is a vector  $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}_{\geq 0}^n$ , often times also expressed as the formal linear combination  $\alpha_1 S_1 + \dots + \alpha_n S_n$ . In this context, the elements of  $\mathcal{S}$  are referred to as the *species* constituting the complex. A *reaction* over a set of complexes  $\mathcal{C}$  is an object of the form  $y \rightarrow y'$  or  $y \rightleftharpoons y'$  for some  $y, y' \in \mathcal{C}$ ,  $y \neq y'$ . The former are referred to as *irreversible* reactions, while the latter are called *reversible*. In either case,  $y$  is called the *reactant* of the reaction, and  $y'$  the *product*.

A *reaction network* is an ordered triple  $G = (\mathcal{S}, \mathcal{C}, \mathcal{R})$  where  $\mathcal{C}$  is a set of complexes over a nonempty, finite set of species  $\mathcal{S} = \{S_1, \dots, S_n\}$ , and  $\mathcal{R} = \{R_1, \dots, R_m\}$  is a nonempty, finite set of reactions over  $\mathcal{C}$ . We write  $\mathcal{R} = \mathcal{R}_{\rightarrow} \cup \mathcal{R}_{\leftrightarrow}$ , where  $\mathcal{R}_{\rightarrow}$  and  $\mathcal{R}_{\leftrightarrow}$  are the disjoint subsets of irreversible and reversible reactions, respectively. We further assume that, for each  $i \in [n]$ , there exists an  $(\alpha_1, \dots, \alpha_n) \in \mathcal{C}$  such that  $\alpha_i > 0$ , and, for each  $y \in \mathcal{C}$ , there exists a reaction in  $\mathcal{R}$  having  $y$  as a reactant or product; in other words,  $\mathcal{S}$  (respectively,  $\mathcal{C}$ ) is the minimal set over which  $\mathcal{C}$  (respectively,  $\mathcal{R}$ ) may be defined. We also assume that

$$y \rightarrow y' \in \mathcal{R} \quad \Rightarrow \quad y \rightleftharpoons y', y' \rightleftharpoons y \notin \mathcal{R}, \quad (3)$$

and

$$y \rightleftharpoons y' \in \mathcal{R} \quad \Rightarrow \quad y' \rightleftharpoons y \notin \mathcal{R}. \quad (4)$$

For each  $j \in [m]$ , let  $\alpha_{1j} S_1 + \dots + \alpha_{nj} S_n$  be the reactant and  $\alpha'_{1j} S_1 + \dots + \alpha'_{nj} S_j$  be the product of reaction  $R_j$ . With this notation, we may define the

$n \times m$  matrix  $N$ ,

$$N_{ij} := \alpha'_{ij} - \alpha_{ij}, \quad i = 1, \dots, n, \quad j = 1, \dots, m,$$

known as the *stoichiometric matrix* of the network. The column-space of  $N$ , which is a subset of  $\mathbb{R}^n$ , is called the *stoichiometric subspace of  $G$* , and denoted by  $\Gamma$ . A vector  $c \in \mathbb{R}^n$  is said to be a *conservation law* of  $G$  if  $c \in \Gamma^\perp$ . The subsets  $(s_0 + \Gamma) \cap \mathbb{R}_{\geq 0}^n$ ,  $s_0 \in \mathbb{R}_{\geq 0}^n$ , are known as the *stoichiometric compatibility classes* of  $G$ .

The system of ordinary differential equations modeling the evolution of the concentrations of the species of the network  $G$  is then given by

$$\frac{ds}{dt} = Nr(s(t)), \quad t \in \mathbb{R}_{\geq 0}, \quad s \in \mathbb{R}_{\geq 0}^n, \quad (5)$$

where  $r = (r_1, \dots, r_m): \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}^m$  is a vector-valued function modeling the kinetic rates of each reaction as functions of the concentrations of the involved species. Unless otherwise noted, we further assume that  $G$  and  $r$  satisfy the hypotheses below.

(G1) There are no *auto-catalytic* reactions, meaning that no species can appear as both reactant and product in any reaction. Thus,  $\alpha_{ij}\alpha'_{ij} = 0$  for any reaction  $R_j \in \mathcal{R}$  and any species  $S_i \in \mathcal{S}$ .

(G2) Each species in  $\mathcal{S}$  takes part in at most two reactions in  $\mathcal{R}$ .

(G3) The network is *conservative*, that is, it has a conservation law  $c \in \mathbb{R}_{> 0}^n$ .

(r1) For each  $j \in [m]$ , if  $R_j$  is irreversible, then  $r_j(s) \geq 0$ ,  $s \in \mathbb{R}_{\geq 0}^n$ ; if  $R_j$  is reversible, then  $r_j = r_j^f - r_j^b$ , where  $r_j^f(s), r_j^b(s) \geq 0$ ,  $s \in \mathbb{R}_{\geq 0}^n$ . Furthermore, all the  $r_j, r_j^f, r_j^b: \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}_{\geq 0}$  have continuously differentiable extensions to a neighborhood  $\mathcal{O}$  of  $\mathbb{R}_{\geq 0}^n$ .

(r2) For each  $j \in [m]$ , and for each  $s = (s_1, \dots, s_n) \in \mathbb{R}_{\geq 0}^n$ ,

(i) if  $R_j$  is irreversible, then

$$s_k = 0 \text{ for some } k \in \{i \in [n] \mid \alpha_{ij} > 0\} \quad \Rightarrow \quad r_j(s) = 0;$$

(ii) if  $R_j$  is reversible, then

$$s_k = 0 \text{ for some } k \in \{i \in [n] \mid \alpha_{ij} > 0\} \quad \Rightarrow \quad r_j^f(s) \geq 0,$$

and

$$s_k = 0 \text{ for some } k \in \{i \in [n] \mid \alpha'_{ij} > 0\} \quad \Rightarrow \quad r_j^b(s) \geq 0.$$

(r3) For each  $j \in [m]$ ,

(i) if  $R_j$  is irreversible, then

$$\frac{\partial r_j}{\partial s_i}(s) \begin{cases} \geq 0, & \text{if } \alpha_{ij} > 0 \\ = 0, & \text{if } \alpha_{ij} = 0. \end{cases}$$

(ii) if  $R_j$  is reversible, then

$$\frac{\partial r_j}{\partial s_i}(s) \begin{cases} \geq 0, & \text{if } \alpha_{ij} > 0 \\ = 0, & \text{if } \alpha_{ij} = 0 \\ \leq 0, & \text{if } \alpha'_{ij} > 0. \end{cases}$$

Furthermore, the inequalities are strict in  $\mathbb{R}_{>0}^n$ .

*Remark 1.* In the literature, one typically defines reaction networks directly from their *reaction graphs* [11], keeping reciprocal reactions as distinct reactions, or, alternatively, by collapsing each pair of reciprocal reactions into a single reversible reaction [3]. Our approach accommodates both extremes, plus anything in between, since it does not preclude the possibility that both  $y \rightarrow y', y' \rightarrow y \in \mathcal{R}$ . In other words, one has the freedom to choose at the beginning which pairs of reciprocal reactions to collapse into a single reversible reaction, and which ones not to.

Hypotheses (r1)–(r3) are satisfied under the most common kinetic assumptions in the literature, namely, mass-action, or more general power-law kinetics, Michaelis-Menten kinetics, or Hill kinetics, as well as combinations of these [3, pages 585–586]. Hypotheses (G1)–(G3) are needed in the context of the graphical conditions for generic/global convergence to equilibria introduced in [3], which we review in the next subsection.

It follows from (r2) and [14, Theorem 5.6] that  $\mathbb{R}_{\geq 0}^n$  is forward invariant for the flow of (5). We then conclude that the interior,  $\mathbb{R}_{>0}^n$ , is also forward-invariant via [1, Remark 16.3(h)]. (See also [16, Section VII].) And in view of (G3), the trajectories of (5) are defined for all positive time, and also precompact.  $\square$

## 2.2 Graph Conditions for Generic Convergence

We now review the concepts and results from [3] that we will need, pointing out that they still hold in our slightly more general setting.

The *directed SR-graph* of a reaction network  $G$  is the directed, bipartite, labeled graph  $G_{SR}^{\rightarrow} = (V_{SR}^{\rightarrow}, E_{SR}^{\rightarrow}, L_{SR}^{\rightarrow})$  defined as follows. The set of vertices  $V_{SR}^{\rightarrow}$  is the disjoint union

$$V_{SR}^{\rightarrow} := \mathcal{S} \cup \mathcal{R} = \mathcal{S} \cup (\mathcal{R}_{\rightarrow} \cup \mathcal{R}_{\leftrightarrow}).$$

The set of edges  $E_{SR}^{\rightarrow}$  and the labeling  $L_{SR}^{\rightarrow}$  are then characterized as follows.

- (i) An ordered pair  $(S_i, R_j) \in \mathcal{S} \times \mathcal{R}_{\rightarrow}$  belongs to  $E_{SR}^{\rightarrow}$  if, and only if  $S_i$  is a reactant of  $R_j$ , that is, if, and only if  $\alpha_{ij} > 0$ .

- (ii) An ordered pair  $(S_i, R_j) \in \mathcal{S} \times \mathcal{R}_{\leftrightarrow}$  belongs to  $E_{SR}^{\rightarrow}$  if, and only if  $S_i$  appears on either side of  $R_j$ , that is, if, and only if  $\alpha_{ij} + \alpha'_{ij} > 0$ .
- (iii) An ordered pair  $(R_j, S_i) \in \mathcal{R} \times \mathcal{S}$  belongs to  $E_{SR}^{\rightarrow}$  if, and only if  $S_i$  is part of  $R_j$  as either a reactant or a product, that is, if, and only if  $\alpha_{ij} + \alpha'_{ij} > 0$ .
- (iv)  $L_{SR}^{\rightarrow}(S_i, R_j) := -\text{sign } N_{ij}$  for every  $(S_i, R_j) \in E_{SR}^{\rightarrow}$ , and  $L_{SR}^{\rightarrow}(R_j, S_i) := -\text{sign } N_{ij}$  for every  $(R_j, S_i) \in E_{SR}^{\rightarrow}$ .

The directed SR-graph is said to be *R-strongly connected* if, for every  $R_j, R_k \in \mathcal{R}$ , there exists a directed path in  $G_{SR}^{\rightarrow}$  connecting  $R_j$  to  $R_k$ .

*Remark 2.* If  $(S_i, R_j), (R_j, S_i) \in E_{SR}^{\rightarrow}$  for some  $i \in [n]$  and some  $j \in [m]$ , then both edges get the same label.  $\square$

The *SR-graph* of  $G$  is the undirected, labeled graph  $G_{SR} = (V_{SR}, E_{SR}, L_{SR})$  where

$$V_{SR} := V_{SR}^{\rightarrow} = \mathcal{S} \cup \mathcal{R},$$

$$E_{SR} := \{\{S_i, R_j\} \mid (S_i, R_j) \in E_{SR}^{\rightarrow} \text{ or } (R_j, S_i) \in E_{SR}^{\rightarrow}\} = \{\{S_i, R_j\} \mid N_{ij} \neq 0\},$$

and

$$L_{SR}(\{S_i, R_j\}) := -\text{sign } N_{ij}, \quad \{S_i, R_j\} \in E_{SR}.$$

In view of Remark 2, the SR-graph is simply the undirected graph underlying the directed SR-graph, in other words, there are no multiple edges connecting any two vertices.

The *R-graph* is the undirected, labeled graph  $G_R = (V_R, E_R, L_R)$  constructed as follows. The vertices set is defined as

$$V_R := \mathcal{R}.$$

Furthermore,

$$E_R := \{\{R_j, R_k\} \mid j \neq k \text{ and } N_{ij}N_{ik} \neq 0 \text{ for some } i \in [n]\},$$

and, for each  $\{R_j, R_k\} \in E_R$ ,

$$L_R(\{R_j, R_k\}) := \{-\text{sign } N_{ij}N_{ik} \mid N_{ij}N_{ik} \neq 0 \text{ and } i \in [n]\}.$$

We emphasize that  $L_R$  is a set-valued function.

The R-graph is said to have the *positive loop property* if every labeled simple loop

$$R_{j_1} \xrightarrow{L_1} R_{j_2} \xrightarrow{L_2} \dots \xrightarrow{L_{\ell-1}} R_{j_\ell} \xrightarrow{L_\ell} R_{j_1}$$

in  $G_R$  has an even number of negative labels, that is,  $L_1 L_2 \cdots L_\ell = 1$  for any choice of  $L_1 \in L_R(\{R_{j_1}, R_{j_2}\})$ ,  $L_2 \in L_R(\{R_{j_2}, R_{j_3}\})$ ,  $\dots$ ,  $L_\ell \in L_R(\{R_{j_\ell}, R_{j_1}\})$ .

The R-graph can be obtained from the SR-graph by placing an edge between two reaction vertices in the R-graph whenever there is a length-2 path connecting the two corresponding reaction vertices in the SR-graph, and labeling that edge with the opposite of the product of the labels along the length-2 path in the



SR-graph. If there are more than one length-2 path connecting any two reaction vertices, it is possible that the corresponding edge in the R-graph gets multiple labels. The positive loop property of the R-graph can then be checked by inspecting the SR-graph from which it is built. This was done in [3], and we cite the relevant result here for ease of reference.

**Definition 1** (E- and O-Loops). Let  $\Lambda: V_0 - V_1 - \dots - V_{2\lambda} - V_0$ ,  $\lambda \in \mathbb{Z}_{>0}$ , be any simple loop in the SR-graph. If

$$\prod_{k=1}^{2\lambda} L_{SR}(\{V_{k-1}, V_k\}) = (-1)^\lambda,$$

then we call  $\Lambda$  an *e-loop*. Otherwise we call it an *o-loop*.  $\triangle$

In the above definition, we know the length of a simple loop in the SR-graph is always an even number because the SR-graph is a bipartite graph.

**Proposition 1.** *Let  $G$  be a reaction network satisfying (G1)–(G2). Then the R-graph has the positive loop property if, and only if all simple loops in the SR-graph are e-loops.*

*Proof.* See [3, Proposition 4.5].  $\square$

*Remark 3.* When the R-graph has the positive loop property, we may associate an *orthant cone*

$$K = \{(x_1, \dots, x_m) \in \mathbb{R}^m \mid \sigma_1 x_1, \dots, \sigma_m x_m \geq 0\}$$

with it by defining the *sign pattern*  $\sigma = (\sigma_1, \dots, \sigma_m) \in \{\pm 1\}^m$  as follows. First suppose the R-graph is connected. Set  $\sigma_1 := 1$ . For  $i \in [m] \setminus \{1\}$ , consider any simple path  $1 = i_0 - i_1 - \dots - i_k = i$  joining 1 and  $i$ , then set

$$\sigma_i := \prod_{d=1}^k L_R(\{R_{i_{d-1}}, R_{i_d}\}). \quad (6)$$

In view of the positive loop property, this definition does not depend on the choice of the path. Indeed, the union of the edges of any two simple paths joining 1 and  $i$  is a union of simple loops. The product of the labels of the edges of the two paths is thus 1, hence the products of the labels of the edges of each of the two paths agree.

If  $G_R$  is not connected, then we apply the procedure to each connected component, starting by setting  $\sigma_i := 1$  for the smallest index  $i \in [m]$  such that  $R_i$  belongs to that component.  $\square$

In what follows, given a reaction network  $G$  such that its R-graph has the positive loop property, we will always assume that  $\sigma = (\sigma_1, \dots, \sigma_m)$  is the sign pattern defined above, and  $K$  the corresponding orthant cone.

**Proposition 2.** *Let  $G$  be a reaction network satisfying (G1)–(G2). Suppose that the  $R$ -graph has the positive loop property, and the directed  $SR$ -graph is  $R$ -strongly connected. Let  $N$  be the stoichiometric matrix, and  $K$  be the orthant cone given by the construction in Remark 3. Then, either*

$$(P1) \quad \ker N \cap K = \{0\},$$

or

$$(P2) \quad \ker N \cap \text{int } K \neq \emptyset.$$

*Proof.* See [3, Lemma 6.1]. □

Recall that the flow of (5) is said to be *bounded-persistent* if  $\omega(s_0) \cap \partial \mathbb{R}_{\geq 0}^n = \emptyset$  for each  $s_0 \in \mathbb{R}_{> 0}^n$ , where

$$\omega(s_0) := \bigcap_{\tau \geq 0} \overline{\bigcup_{t \geq \tau} \{\sigma(t, s_0)\}}$$

is the *omega-limit set* of  $s_0$ .

**Proposition 3.** *Let  $G$  be a reaction network satisfying (G1)–(G3) and (r1)–(r3). Suppose that the flow of (5) is bounded-persistent. Suppose, in addition, that the  $R$ -graph has the positive loop property and that the directed  $SR$ -graph is  $R$ -strongly connected. Then,*

- (i) *if (P1) holds, then there exists a Lebesgue measure-zero  $D \subseteq \mathbb{R}_{> 0}^n$  such that all solutions of (5) starting in  $\mathbb{R}_{> 0}^n \setminus D$  converge to the set of equilibria, and*
- (ii) *if (P2) holds, then all solutions of (5) starting in  $\mathbb{R}_{> 0}^n$  converge to an equilibrium. Furthermore, this equilibrium is unique within each stoichiometric compatibility class.*

*Proof.* See [3, Corollary 1 and Theorem 2]. The cone constructed from the  $R$ -graph in Remark 3 is the same as the cone given by [3, Corollary 1]. □

### 3 Main Results

This work is essentially about how the graphical conditions for generic/global convergence to equilibria reviewed in Proposition 3 are invariant under the removal of so-called intermediates. However, the removal of intermediates in the sense they are typically defined in the reaction network literature [6] often gives rise to auto-catalytic reactions, which are not allowed in our formalism because of (G1). As we will see, if the problematic species were to appear only in the reactions emerging from the removal of intermediates, and with the same stoichiometric coefficients in both reactant and product sides, then they could be simply “cancelled out,” so that the network obtained from their removal still satisfies (G1).

We begin this section by giving a formal description of the procedure of removal of intermediates. We state our main results in Subsection 3.2, and discuss several examples from the literature in Subsection 3.3. In Subsection 3.4, we briefly contrast our working definition of intermediates with other variants in the literature, giving some examples and counterexamples motivating our choices in this context.

### 3.1 Removal of Intermediates

Let  $G = (\mathcal{S}, \mathcal{C}, \mathcal{R})$  be a reaction network. For each  $y = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}_{\geq 0}^m$ , we denote

$$\text{supp } y := \{S_i \in \mathcal{S} \mid \alpha_i > 0\}.$$

Given a  $Y \in \mathcal{S}$ , consider the following two properties.

- (I1)  $Y \in \mathcal{C}$ , and  $\text{supp } y \cap \text{supp } Y = \emptyset$  for every complex  $y \in \mathcal{C} \setminus \{Y\}$ .
- (I2) There exist **unique**  $y = \alpha_1 S_1 + \dots + \alpha_n S_n$  and  $y' = \alpha'_1 S_1 + \dots + \alpha'_n S_n$  in  $\mathcal{C} \setminus \{Y\}$ ,  $y \neq y'$ , such that
  - (i) either  $y \rightarrow Y$  or  $y \rightleftharpoons Y$  is a reaction in  $\mathcal{R}$ ,
  - (ii) either  $Y \rightarrow y'$  or  $Y \rightleftharpoons y'$  is a reaction in  $\mathcal{R}$ ,
  - (iii)  $\sum_{S_i \in \mathcal{E}} \alpha_i S_i = \sum_{S_i \in \mathcal{E}} \alpha'_i S_i =: e$ , where  $\mathcal{E} := \text{supp } y \cap \text{supp } y'$ , and
  - (iv)  $y - e \rightarrow y' - e$ ,  $y' - e \rightarrow y - e$ ,  $y - e \rightleftharpoons y' - e$ , and  $y' - e \rightleftharpoons y - e$  are not reactions in  $\mathcal{R}$ .

If (I1) and (I2) hold, then we may construct a reaction network  $G^* = (\mathcal{S}^*, \mathcal{C}^*, \mathcal{R}^*)$  as follows. We define  $\mathcal{R}^* := \mathcal{R}_c^* \cup \mathcal{R}_Y^*$ , where  $\mathcal{R}_c^*$  is identified with the subset of  $\mathcal{R}$  of reactions not involving the complex  $Y$ , and

$$\mathcal{R}_Y^* := \begin{cases} \{y - e \rightleftharpoons y' - e\}, & \text{if } y \rightleftharpoons Y, Y \rightleftharpoons y' \in \mathcal{R} \\ \{y - e \rightarrow y' - e\}, & \text{if } y \rightarrow Y \in \mathcal{R} \text{ or } Y \rightarrow y' \in \mathcal{R}. \end{cases}$$

We set  $\mathcal{C}^*$  to be the set of reactant and product complexes in the reactions in  $\mathcal{R}^*$ , and  $\mathcal{S}^*$  is set of species that are part of some complex in  $\mathcal{C}^*$ . In the above description, we think of the reactant and product sides of a reaction in  $\mathcal{R}$  or  $\mathcal{R}^*$  as the formal linear combinations of participating species.

**Definition 2** (Intermediates). Let  $G = (\mathcal{S}, \mathcal{C}, \mathcal{R})$  be a reaction network. We say that  $Y \in \mathcal{S}$  is an *intermediate* if (I1) and (I2) hold. In this case, the reaction network  $G^* = (\mathcal{S}^*, \mathcal{C}^*, \mathcal{R}^*)$  constructed as above is called the *reduction of  $G$  by the removal of the intermediate  $Y$* .  $\triangle$

*Remark 4.* Note that  $\text{supp } e$  might be empty. If  $\text{supp } e$  is not empty, then, in view of (G2), no species in it can take part in any other reaction in  $G$  besides  $y - Y$  and  $Y - y'$ , where the dash ‘—’ is a placeholder for ‘ $\rightarrow$ ’ or ‘ $\rightleftharpoons$ .’ In particular, no species in  $\text{supp } e$  is present in  $G^*$ .  $\square$

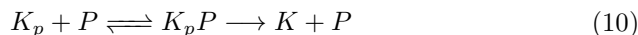
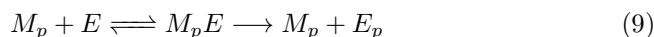
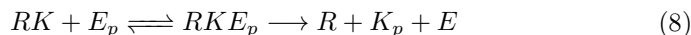
*Remark 5.* Recall that we are assuming reaction networks to satisfy (G1)–(G3). We note that the removal of an intermediate does not break any of these properties. Indeed,  $\text{supp}(y - e) \cap \text{supp}(y' - e) = \emptyset$  by construction, so  $G^*$  satisfies (G1) whenever  $G$  does. Furthermore, it follows directly from the construction that no species in  $G^*$  can take part in more than two reactions, that is, (G2) also holds for  $G^*$ , as long as it already did for  $G$ . Finally, it follows from [11, Theorems 1(*iv*) and 2(*iv*)] that (G3) is preserved by the removal of an intermediate in the sense of Definition 2, which can be seen as a special case of the iterative removal of sets of intermediates or catalysts in the sense of [11]. We omit the details. (See Remark 1 also.)  $\square$

Given  $Y_1, \dots, Y_p \in \mathcal{S}$ , set  $G_p = (\mathcal{S}_p, \mathcal{C}_p, \mathcal{R}_p) := G$ , and suppose that, for  $j = p, \dots, 1$ , we recursively have  $Y_1, \dots, Y_j \in \mathcal{S}_j$ , that the species  $Y_j$  is an intermediate of  $G_j$ , and then define  $G_{j-1} = (\mathcal{S}_{j-1}, \mathcal{C}_{j-1}, \mathcal{R}_{j-1})$  to be the reaction network obtained from  $G_j$  by the removal of the intermediate  $Y_j$ .

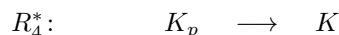
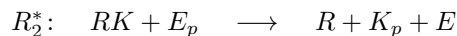
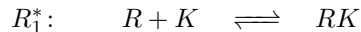
**Definition 3** (Successive Removal of Intermediates). The reaction network  $G_0$  obtained in the above construction is referred to as the *reduction of  $G$  by the successive removal of intermediates  $Y_p, \dots, Y_1$* .  $\triangle$

*Remark 6.* Observe that  $G_0$  does not depend on the order in which the intermediates are removed. This can be shown as in the proof of [11, Theorem 3] (see Remark 5 also). We omit the details.  $\square$

*Example 1* (The RKIP Network). Consider the *RKIP network* discussed in [3, Example 2], displayed below in slightly modified notation as equations (7)–(10).



The reaction network obtained by the removal of the intermediate  $M_pE$  consists of (7), (8), (10), plus the reaction  $E \longrightarrow E_p$  (the species  $M_p$  is cancelled out upon the removal of  $M_pE$ ). We may further remove the intermediate  $RKE_p$ , then the intermediate  $K_pP$  (leading to  $P$  being also cancelled out), eventually obtaining



as the reduced reaction network.  $\diamond$

*Example 2* (Single-Phosphorylation Mechanism). Consider the one-site phosphorylation mechanism (1) discussed in the introduction. The reaction network obtained by the successive removal of intermediates  $S_0E$  and  $S_1F$  is given by



We emphasize that, in our formalism, the reduced network consists of the two reactions  $R_1$  and  $R_2$ , and not of the single reversible reaction  $S_0 \rightleftharpoons S_1$ .  $\diamond$

### 3.2 Invariance under the Removal of an Intermediate

**Theorem 1.** *Let  $G$  be a reaction network satisfying (G1)–(G2). Suppose  $G^*$  is a reaction network obtained from  $G$  by the successive removal of intermediates. Then  $G^*$  also satisfies (G1)–(G2) and, furthermore,*

- (i) *the directed SR-graph of  $G^*$  is R-strongly connected if, and only if the directed SR-graph of  $G$  is R-strongly connected, and*
- (ii) *the R-graph of  $G^*$  has the positive loop property if, and only if the R-graph of  $G$  has the positive loop property.*

Furthermore, if these two graphical conditions are met, then (P1) and (P2) are also invariant under the removal of sets of intermediates.

**Theorem 2.** *Let  $G$  be a reaction network satisfying (G1)–(G2). Suppose  $G^*$  is a reaction network obtained from  $G$  by the successive removal of intermediates. Suppose, in addition, that the R-graph of  $G^*$  has the positive loop property, and the directed SR-graph of  $G^*$  is R-strongly connected. Let  $N$  and  $N^*$  be the stoichiometric matrices of  $G$  and  $G^*$ , and  $K$  and  $K^*$  the orthant cones constructed in Remark 3 from the R-graphs of  $G$  and  $G^*$ , respectively. Then,*

$$\ker N \cap K = \{0\} \quad \Leftrightarrow \quad \ker N^* \cap K^* = \{0\},$$

and

$$\ker N \cap \text{int } K \neq \emptyset \quad \Leftrightarrow \quad \ker N^* \cap \text{int } K^* \neq \emptyset.$$

In view of Theorem 1, the graphical hypotheses on the directed SR- and R-graphs in Proposition 3 for  $G$  can be checked in  $G^*$ . And in view of Theorem 2, if these hypotheses are satisfied, then (P1)/(P2) can also be checked in  $G^*$ . The hypothesis of bounded-persistence in Proposition 3 can be checked using the graphical conditions in [2]. As shown in [11], these graphical conditions for bounded-persistence can also be checked in  $G^*$ . We have thus devised a method to study the qualitative property of generic or global convergence of a reaction network by analyzing a reduced network associated with it.

The proofs of Theorems 1 and 2 will be given in Section 4. We first illustrate the results with a few examples.

### 3.3 Examples

*Example 3 (The RKIP Network).* Consider the RKIP network discussed in Example 1, which was reduced by the successive removal of intermediates  $M_pE$ ,  $RKE_p$  and  $K_pP$ . The R-graph of the reduced network is shown in Figure 3b. One can readily see that it has the positive loop property. The directed SR-graph of the reduced network is shown in Figure 3a. One can also readily see

that it is R-strongly connected. We conclude via Theorem 1 that the directed SR- and the R-graphs of the original RKIP network have the same properties.

The R-graph of the reduced network yields, via Remark 3, the orthant cone  $K^* := \mathbb{R}_{\geq 0}^4$ . Furthermore, each of the species  $R, K, RK, E_p, E, K_p$  appears in exactly two reactions, once as a reactant, once as a product, both times with stoichiometric coefficient 1. Therefore,  $(1, 1, 1, 1) \in \ker N^*$ , showing that  $\ker N^* \cap \text{int } K^* \neq \emptyset$ . It follows from Theorem 2 that  $\ker N \cap \text{int } K \neq \emptyset$ , where  $N$  is the stoichiometric matrix of the original RKIP network, and  $K$  is the orthant cone obtained for its R-graph via Remark 3.

The property of bounded-persistence for the flow of  $G$  can also be checked directly on  $G^*$  (see [11, Theorems 1 and 2] and Remark 1).

We conclude via Proposition 3 that, under kinetic assumptions (r1)–(r3), the RKIP network from Example 1 has that property that each stoichiometric compatibility class has a unique equilibrium to which all trajectories starting with strictly positive concentrations converge.  $\diamond$

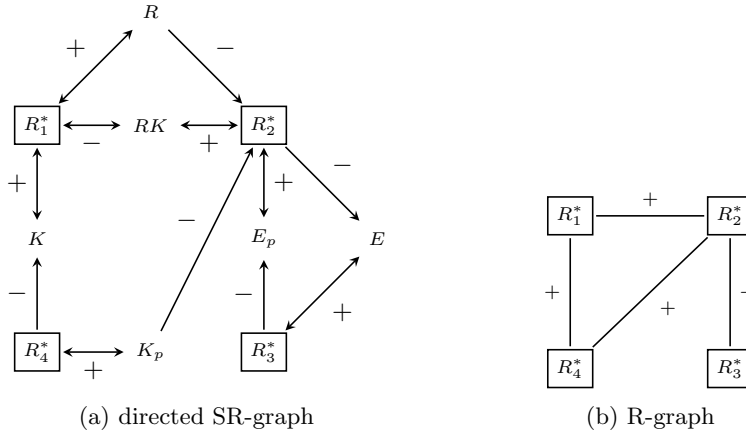
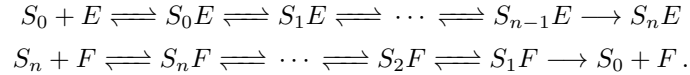


Figure 3: The directed SR- and the R-graph of the reduced RKIP network from Example 1.

*Example 4* (Processive  $n$ -Site Phosphorylation Mechanism). Consider the sequential and processive  $n$ -site phosphorylation mechanism described by the reaction network



(See [4], and references therein.) Note that the one-site mechanism from Example 2 is the special case when  $n = 1$  of this mechanism. The reaction network obtained by the successive removal of the intermediates  $S_0E, \dots, S_{n-1}E, S_nF, \dots, S_1F$  is given by



The R-graph of the reduced network (Figure 4a) has no loops, so, it vacuously has the positive loop property. Furthermore, the directed SR-graph of the reduced network (Figure 4b) is R-strongly connected. It follows from Theorem 1 that the directed SR- and the R-graphs of the original  $n$ -site phosphorylation network have the same properties.

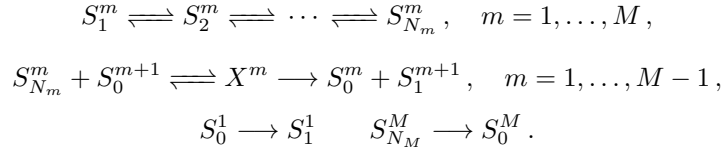
The R-graph of the reduced network yields, via Remark 3, the orthant cone  $K^* := \mathbb{R}_{\geq 0}^2$ . One can argue as in Example 3 that  $(1, 1) \in \ker N^*$ , showing that  $\ker N^* \cap \text{int } K^* \neq \emptyset$ , and so  $\ker N \cap \text{int } K \neq \emptyset$  via Theorem 2. Finally, one can once again show that the flow of  $G$  is bounded-persistent via [11, Theorems 1 and 2] and Remark 1.

It follows from Proposition 3 that, under kinetic assumptions (r1)–(r3), each stoichiometric compatibility class has a unique equilibrium to which all trajectories starting with strictly positive concentrations converge.  $\diamond$



Figure 4: The directed SR- and the R-graph of the reduced  $n$ -site phosphorylation network from Example 4.

*Example 5* (A Phosphorelay). Consider the general phosphorelay system studied in [10]. The underlying reaction network consists of the reactions



For each  $m \in [M]$  and each  $n \in [N_m]$ ,  $S_n^m$  represents the  $m^{\text{th}}$  substrate (out of  $M \geq 1$ ), phosphorylated at its  $n^{\text{th}}$  site (out of  $N_m \geq 1$ ), and  $S_0^m$  corresponds to the unphosphorylated state of the  $m^{\text{th}}$  substrate. The phosphate group can be transferred sequentially from site to site within the same substrate, or via the formation of an intermediate complex  $X^m$  from  $S_{N_m}^m$  to  $S_1^{m+1}$ ,  $m = 1, \dots, M-1$ . The methods in [3] were employed in [10] to show that, under mass-action kinetics, the phosphorelay has a unique nonnegative equilibrium to which all solutions starting with positive concentrations converge.

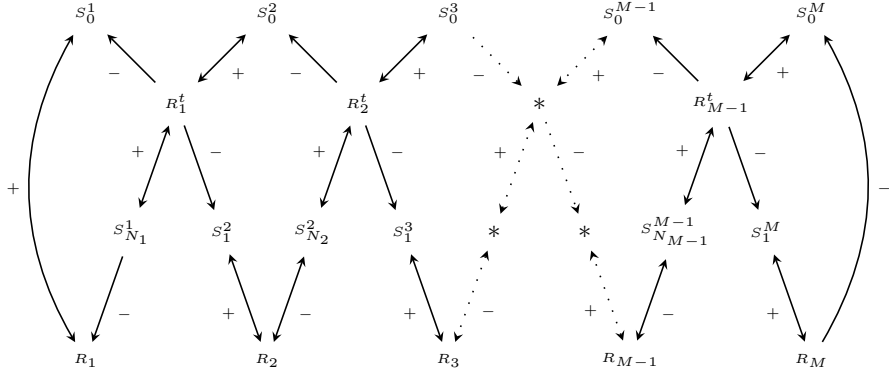
First note that each species in the phosphorelay takes part in exactly two reactions. Thus, (G2) is satisfied. Now

$$\{S_1^1, S_2^1, \dots, S_{N_1-1}^1, S_2^2, \dots, S_{N_2-1}^2, \dots, S_2^M, \dots, S_{N_M-1}^M, S_{N_M}^M, X^1, \dots, X^{M-1}\}$$

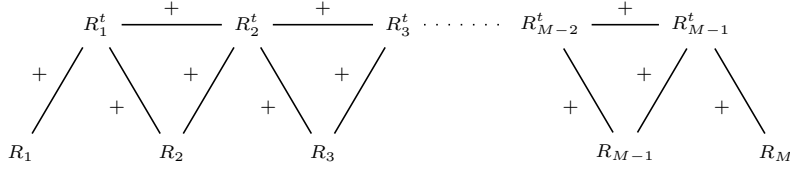
is a set of intermediates. The network obtained after their removal is given by

$$\begin{array}{llll}
R_1: & S_0^1 \longrightarrow S_{N_1}^1 & & \\
R_2: & S_1^2 \rightleftharpoons S_{N_2}^2 & R_1^t: & S_{N_1}^1 + S_0^2 \longrightarrow S_0^1 + S_1^2 \\
& \vdots & & \vdots \\
R_{M-1}: & S_1^{M-1} \rightleftharpoons S_{N_{M-1}}^{M-1} & R_{M-1}^t: & S_{N_{M-1}}^{M-1} + S_0^M \longrightarrow S_0^{M-1} + S_1^M . \\
R_M: & S_1^M \longrightarrow S_0^M & & 
\end{array}$$

The directed SR-graph and the R-graph of the reduction are sketched in Figures 5a and 5b, where they can be readily seen to have, respectively, the R-strong connectedness and positive loop properties. It follows from Theorem 1 that the directed SR-graph and the R-graph of the original network have, respectively, the R-strong connectedness and positive loop properties.



(a) directed SR-graph



(b) R-graph

Figure 5: The SR- and the R-graph of the reduction of the phosphorelay system.

The R-graph of the reduced network (Figure 5b) yields, via Remark 3, the orthant cone  $K^* := \mathbb{R}_{\geq 0}^{2M-1}$ . The same argument as in the previous two examples shows that  $(1, \dots, 1) \in \mathbb{R}^{2M-1}$  belongs to the kernel of the stoichiometric matrix  $N^*$  of the reduced network. Thus,  $\ker N^* \cap \text{int } K^* \neq \emptyset$ , and so  $\ker N \cap \text{int } K \neq \emptyset$  by Theorem 2. As before, bounded-persistent follows via [11, Theorems 1 and 2] and Remark 1.

We conclude via Proposition 3 that, within each stoichiometric compatibil-



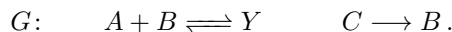
ity class, there exists a unique nonnegative equilibrium to which all solutions starting with strictly positive concentrations converge.  $\diamond$

### 3.4 Further Comments on Definition 2

Our working definition of intermediates in this paper is somewhat more restrictive than in [11], where a related study of invariance of qualitative properties of reaction networks under the removal of sets of intermediates and catalysts was carried out. We conclude this section with a discussion of the differences, and show some examples of what may go wrong with the kinds of intermediates precluded in our working definition.

#### The $y \rightleftharpoons Y \in \mathcal{R}$ Case

In [11], a species  $Y$  would still be considered an intermediate if  $y = y'$  in (I2). In this case,  $G^*$  is defined by simply removing the reaction  $y \rightleftharpoons Y$  from  $G$  (or removing the reactions  $y \rightarrow Y$  and  $Y \rightarrow y$ , if that is the case). The R-strong connectedness property of the directed SR-graph is not invariant under the removal of intermediates of this type. To see this, consider the reaction network



The reaction network obtained by removing  $Y$  as described above is

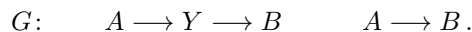


The directed SR-graph of  $G^*$  is R-strongly connected, while that of  $G$  is not.

#### The $y \rightarrow Y \rightarrow y', (y - e) \rightarrow (y' - e) \in \mathcal{R}$ Case

We first note that, if (G2) holds, and the SR-graph of the network is assumed to be connected, then there cannot be any other reactions in the network besides these three. Indeed, each species constituting  $y$ ,  $y'$  or  $Y$  is already linked in the SR-graph to two reaction nodes. So, neither of these species can be involved in any other reaction. Therefore, precluding this kind of intermediates in our working definition is far less restrictive than it might seem at first — the bulk of the constrain comes from (G2) already.

In (I2)(iv), if we allow for  $(y - e) \rightarrow (y' - e)$  to be a reaction in  $\mathcal{R}$ , then the R-strong connectedness condition for monotonicity might not be invariant under the removal of intermediates of this type. To see this, consider the reaction network

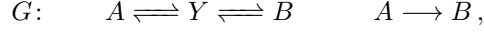


By removing  $Y$ , we get the reaction network

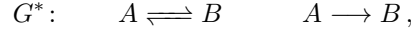


As before, the directed SR-graph of  $G^*$  is R-strongly connected, while that of  $G$  is not.

Furthermore, the reaction network obtained by removing  $Y$  might not be well defined without (I2)(iv). For instance, if



then



which violates (3) and (4).

## 4 Proofs of Theorems 1 and 2

Note that it suffices to prove Theorems 1 and 2 for the removal of a single intermediate  $Y$ . The general result then follows by induction on the number of intermediates successively removed. We have four cases to consider, depending on how  $Y$  appears in  $G$ , all of which are captured by

$$y + \sum_{i=1}^p \gamma_i E_i \text{ --- } Y \text{ --- } y' + \sum_{i=1}^p \gamma_i E_i, \quad (11)$$

where each ‘---’ may mean either ‘ $\longrightarrow$ ’ or ‘ $\rightleftharpoons$ ’,

$$e := \sum_{i=1}^p \gamma_i E_i$$

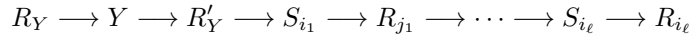
may be an empty sum, and  $\text{supp } y \cap \text{supp } y' = \emptyset$ .

### 4.1 Proof of Theorem 1(i)

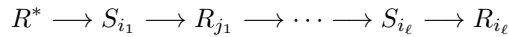
Although the directed SR-graph will not be quite the same in all cases, they can still be treated together. Denote

$$R_Y: y + \sum_{i=1}^p \gamma_i E_i \text{ --- } Y, \quad R'_Y: Y \text{ --- } y' + \sum_{i=1}^p \gamma_i E_i.$$

By reordering the species in  $\mathcal{S} = \{S_1, \dots, S_n\}$ , if necessary, we may write  $y = \alpha_1 S_1 + \dots + \alpha_k S_k$  and  $y' = \alpha'_1 S'_1 + \dots + \alpha'_{k'} S'_{k'}$  for some  $\alpha_1, \dots, \alpha_k, \alpha'_1, \dots, \alpha'_{k'} > 0$  and some  $S_1, \dots, S_k, S'_1, \dots, S'_{k'} \in \mathcal{S}$ . Figure 6a illustrates the directed SR-graph of  $G$ . Note that there is always a directed path  $R_Y \longrightarrow Y \longrightarrow R'_Y$  in the graph. This path is replaced by the reaction node  $R^* = y - y'$  in the directed SR-graph of  $G^*$  (see Figure 6b). It is not difficult to see that  $G_{SR}^{\rightarrow}$  is R-strongly connected if, and only if  $(G^*)_{SR}^{\rightarrow}$  is also R-strongly connected. For instance,



is a directed path in  $G_{SR}^{\rightarrow}$  if, and only if



is a directed path in  $(G^*)_{SR}^{\rightarrow}$ . Similar pairings can be made for paths ending or passing through  $R_Y$ ,  $R'_Y$  or  $R^*$  in their respective graphs, and paths not passing through either of these reaction nodes can be found in a graph if, and only if they can be found in the other.  $\square$

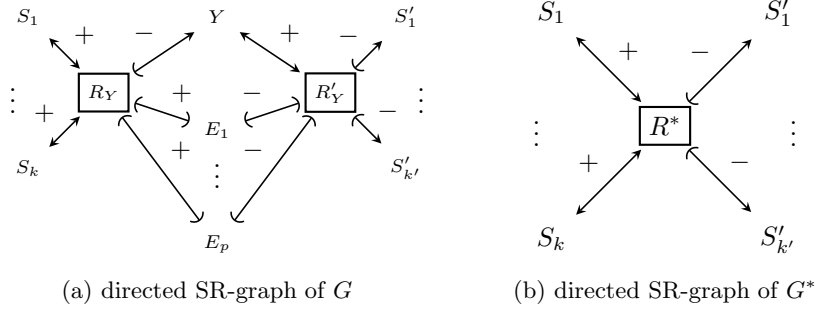


Figure 6: The directed SR-graphs of  $G$  and  $G^*$ . A rounded arrow head at the end of an edge indicates that the edge is only in the graph in that direction in some of the cases in (11). For instance, the edge  $Y \rightarrow R_Y$  is only in the directed SR-graph of  $G$  if the first ‘ $-$ ’ in (11) is a ‘ $\rightleftharpoons$ .’ The (undirected) SR-graphs of  $G$  and  $G^*$  are obtained by simply ignoring the arrow heads altogether.

## 4.2 Proof of Theorem 1(ii)

To study the positive loop property of the R-graph of  $G$  and  $G^*$ , we will use Proposition 1, together with the lemma below.

**Lemma 1.** *A simple loop in the SR-graph is an e-loop if, and only if it contains an even number of segments  $S_i - R_j - S_k$ ,  $S_i, S_k \in \mathcal{S}$ ,  $R_j \in \mathcal{R}$ , such that  $L_{SR}(\{S_i, R_j\}) = L_{SR}(\{R_j, S_k\})$ .*

*Proof.* See [3, Lemma 4.4].  $\square$

Let  $R_Y, R'_Y, R^*, \alpha_1, \dots, \alpha_k, \alpha'_1, \dots, \alpha'_{k'}, S_1, \dots, S_k, S'_1, \dots, S'_{k'}$  be as in the proof of Theorem 1(i).

On the one hand,  $L$  is a simple e-loop (respectively, o-loop) of  $G_{SR}^*$  which does not go through  $R^*$  if, and only if it is a simple e-loop (respectively, o-loop) of  $G_{SR}$  which does not go through either of  $R_Y, Y$  and  $R'_Y$ .

On the other hand, the simple loops of  $G_{SR}^*$  which go through  $R^*$  are in one-to-one correspondence with the simple loops of  $G_{SR}$  going through  $R_Y, Y$  or  $R'_Y$ . This correspondence is established as follows. Any simple loop  $L$  in  $G_{SR}^*$  which goes through  $R^*$  has the form

$$R^* - S_{out} - R_{i_1} - S_{i_1} - \dots - S_{i_\ell} - R_{i_\ell} - S_{in} - R^*$$

for some pairwise distinct  $R_{i_1}, \dots, R_{i_\ell} \in \mathcal{R}^* \cap \mathcal{R}$  and  $S_{out}, S_{i_1}, \dots, S_{i_\ell}, S_{in} \in \mathcal{S}^* \cap \mathcal{R}$ . If  $S_{out}, S_{in} \in \{S_1, \dots, S_k\}$  (respectively,  $S_{out}, S_{in} \in \{S'_1, \dots, S'_{k'}\}$ ),

then we need only replace  $R^*$  with  $R_Y$  (respectively,  $R'_Y$ ). If  $S_{out}$  belongs to one of the sets  $\{S_1, \dots, S_k\}$  and  $\{S'_1, \dots, S'_{k'}\}$ , and  $S_{in}$  belongs to the other, then we need only replace  $R^*$  by the segment  $R_Y - Y - R'_Y$ . Now note that this correspondence also takes e-loops (respectively, o-loops) to e-loops (respectively, o-loops). Indeed, this follows from Lemma 1. If  $S_{in}$  and  $S_{out}$  belong to the same set, the simple loop in  $G_{SR}$  has the same number of edges as its corresponding loop in  $G_{SR}^*$ , and they both have the same sign pattern. If  $S_{in}$  and  $S_{out}$  belong to different sets, then the consecutive edges in the segments

$$S_{in} - R^* - S_{out} \quad \text{and} \quad S_{in} - R_Y - Y - R'_Y - S_{out}$$

have opposite signs, as one can see in Figures 6a and 6b, so the number of segments  $S_i - R_j - S_k$  such that  $S_i - R_j$  and  $R_j - S_k$  have the same sign does not change from the path in  $G_{SR}^*$  to the corresponding path in  $G_{SR}$ .

We conclude that every simple loop in  $G_{SR}$  is an e-loop if, and only if every simple loop in  $G_{SR}^*$  is an e-loop. Since (G2) holds, it follows from Proposition 1 that the R-graph of  $G$  has the positive loop property if, and only if the R-graph of  $G^*$  also has the positive loop property.  $\square$

### 4.3 Proof of Theorem 2

Once we understand the relationships between  $\ker N$  and  $\ker N^*$  and between  $K$  and  $K^*$ , the proof of the theorem will follow somewhat effortlessly.

#### Relationship between $\ker N$ and $\ker N^*$

We first consider the case in which  $e$  is nontrivial. By reordering the species and reactions so that  $Y, E_1, \dots, E_p$ , and the reactions  $y - Y$  and  $Y - y'$  appear at the end, if necessary, we may write the stoichiometric matrices  $N$  and  $N^*$  of, respectively,  $G$  and  $G^*$  as

$$N = \left[ \begin{array}{ccc|cc} & & & -\alpha_1 & \alpha'_1 \\ & & & \vdots & \vdots \\ & & & -\alpha_n & \alpha'_n \\ \hline 0 & \cdots & 0 & 1 & -1 \\ \hline 0 & \cdots & 0 & -\gamma_1 & \gamma_1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & -\gamma_p & \gamma_p \end{array} \right] \quad \text{and} \quad N^* = \left[ \begin{array}{c|c} & \begin{bmatrix} \alpha'_1 - \alpha_1 \\ \vdots \\ \alpha'_n - \alpha_n \end{bmatrix} \end{array} \right],$$

for some  $n \times (m-1)$  matrix  $N_c^*$ , where  $n$  is the number of nonintermediate species,  $m-1$  is the number of reactions not involving  $Y$ , and where we write  $y = \alpha_1 S_1 + \cdots + \alpha_n S_n + \gamma_1 E_1 + \cdots + \gamma_p E_p$  and  $y' = \alpha'_1 S_1 + \cdots + \alpha'_n S_n + \gamma_1 E_1 + \cdots + \gamma_p E_p$ . Thus,

$$\ker N = \{(v_1, \dots, v_{m-1}, v_m, v_m) \in \mathbb{R}^{m+1} \mid (v_1, \dots, v_m) \in \ker N^*\}. \quad (12)$$

In case  $p = 0$ , the argument is basically the same. The only difference is that  $N$  does not have the  $p$  bottom-most rows corresponding to the catalysts. The relationship between  $\ker N$  and  $\ker N^*$  is also given by (12).

### Relationship between $K$ and $K^*$

We order the reactions  $R_1, \dots, R_m, R_{m+1}$  of  $G$  and  $R_1^*, \dots, R_m^*$  of  $G^*$  as above, so that  $R_j$  and  $R_j^*$  are identified for  $j = 1, \dots, m-1$ , and  $R_m = y - Y$ ,  $R_{m+1} = Y - y'$ , and  $R_m^* = y - y'$ . Note that the R-graph of  $G^*$  could be obtained from the R-graph of  $G$  by simply collapsing the edge  $\{R_m, R_{m+1}\}$  in  $G_R$  into the vertex  $R_m^*$  in  $G_R^*$  (refer to Figures 6a and 6b). Indeed, we have

$$L(\{R_m, R_{m+1}\}) = 1, \quad (13)$$

and, in view of the positive loop property,  $L(\{R_j, R_m\}) = L(\{R_j, R_{m+1}\})$  for any  $j \in [m-1]$  such that  $\{R_j, R_m\}, \{R_j, R_{m+1}\} \in E_R$ . Thus,

$$L(\{R_i, R_j\}) = L^*(\{R_i^*, R_j^*\}), \quad \forall i, j \in [m] : \{R_i, R_j\} \in E_R, \quad (14)$$

and

$$L(\{R_j, R_{m+1}\}) = L^*(\{R_j^*, R_m^*\}), \quad \forall j \in [m-1] : \{R_j, R_{m+1}\} \in E_R. \quad (15)$$

Let  $\sigma = (\sigma_1, \dots, \sigma_m, \sigma_{m+1})$  and  $\sigma^* = (\sigma_1^*, \dots, \sigma_m^*)$  be the sign patterns of the cones  $K$  and  $K^*$  constructed via Remark 3 for  $G$  and  $G^*$ , respectively. It follows from (6) and (13)–(15) that  $\sigma_m = \sigma_{m+1}$  and  $\sigma_j^* = \sigma_j$ ,  $j = 1, \dots, m-1$ .

### Proof of Theorem 2

We may summarize the discussion above as a lemma.

**Lemma 2.** *Assume the same hypotheses as in Theorem 2. Then*

$$\ker N = \{(v_1, \dots, v_{m-1}, v_m, v_m) \in \mathbb{R}^{m+1} \mid (v_1, \dots, v_m) \in \ker N^*\}$$

and the sign pattern  $\sigma$  of  $K$  is given by

$$\sigma = (\sigma_1^*, \dots, \sigma_m^*, \sigma_m^*),$$

where  $(\sigma_1^*, \dots, \sigma_m^*) = \sigma^*$  is the sign pattern of  $K^*$ .

It follows from the lemma that

$$(v_1, \dots, v_m) \in \ker N^* \cap K^* \quad \Leftrightarrow \quad (v_1, \dots, v_m, v_m) \in \ker N \cap K,$$

and, moreover,

$$(v_1, \dots, v_m) \in \text{int } K^* \quad \Leftrightarrow \quad (v_1, \dots, v_m, v_m) \in \text{int } K.$$

This establishes Theorem 2.

## Acknowledgements

Elisenda Feliu, Michael Marcondes de Freitas and Carsten Wiuf acknowledge funding from the Danish Research Council of Independent Research.

## References

- [1] H. Amann. *Ordinary differential equations*, volume 13 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1990. An introduction to nonlinear analysis, Translated from the German by Gerhard Metzen.
- [2] D. Angeli, P. De Leenheer, and E. D. Sontag. A Petri net approach to the study of persistence in chemical reaction networks. *Math. Biosci.*, 210(2):598–618, 2007.
- [3] D. Angeli, P. De Leenheer, and E. D. Sontag. Graph-theoretic characterizations of monotonicity of chemical networks in reaction coordinates. *J. Math. Biol.*, 61(4):581–616, 2010.
- [4] C. Conradi and A. Shiu. A global convergence result for processive multisite phosphorylation systems. *Bull. Math. Biol.*, 77(1):126–155, 2015.
- [5] A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. Portland Press, London, third edition, 2004.
- [6] E. Feliu and C. Wiuf. Simplifying biochemical models with intermediate species. *J. R. Soc. Interface*, 10(87):20130484, 2013.
- [7] M. W. Hirsch. Stability and convergence in strongly monotone dynamical systems. *J. Reine Angew. Math.*, 383:1–53, 1988.
- [8] M. W. Hirsch and H. Smith. Monotone dynamical systems. In *Handbook of differential equations: ordinary differential equations. Vol. II*, pages 239–357. Elsevier B. V., Amsterdam, 2005.
- [9] E. L. King and C. Altman. A schematic method of deriving the rate laws for enzyme-catalyzed reactions. *J. Phys. Chem.*, 60:1375–1378, 1956.
- [10] M. Knudsen, E. Feliu, and C. Wiuf. Exact analysis of intrinsic qualitative features of phosphorelays using mathematical models. *J. Theoret. Biol.*, 300:7–18, 2012.
- [11] M. Marcondes de Freitas, E. Feliu, and C. Wiuf. Intermediates, catalysts, persistence, and boundary steady states, 2015. Submitted.
- [12] O. Radulescu, A. N. Gorban, A. Zinovyev, and V. Noel. Reduction of dynamical biochemical reactions networks in computational biology. *Frontiers in Genetics*, 3, 2012.

- [13] M. Sáez, C. Wiuf, and E. Feliu. Graphical reduction of reaction networks by linear elimination of species. *J. Math. Biol.*, First online: 24 May, 2016.
- [14] G. V. Smirnov. *Introduction to the theory of differential inclusions*, volume 41 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.
- [15] H. L. Smith. *Monotone dynamical systems*, volume 41 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1995. An introduction to the theory of competitive and cooperative systems.
- [16] E. D. Sontag. Structure and stability of certain chemical networks and applications to the kinetic proofreading model of T-cell receptor signal transduction. *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, 46(7):1028–1047, 2001.
- [17] C. Widmann, G. Spencer, M. B. Jarpe, and G. L. Johnson. Mitogen-activated protein kinase: Conservation of a three-kinase module from yeast to human. *Physiol. Rev.*, 79(1):143–180, 1999.