Graphical reduction of reaction networks by linear elimination of species

Saez Cornellana, Meritxell; Wiuf, Carsten; Feliu, Elisenda

Published in:
Journal of Mathematical Biology

DOI:
10.1007/s00285-016-1028-y

Publication date:
2017

Graphical reduction of reaction networks by linear elimination of species

Meritxell Sáez Carsten Wiuf Elisenda Feliu

September 11, 2015

Abstract

We give a graphically based procedure to reduce a reaction network to a smaller reaction network with fewer species after linear elimination of a set of noninteracting species. We give a description of the reduced reaction network, its kinetics and conservations laws, and explore properties of the network and its kinetics. We conclude by comparing our approach to an older similar approach by Temkin and co-workers. Finally, we apply the procedure to biological examples such as substrate mechanisms, post-translational modification systems and networks with intermediates (transient) steps.

Key words: Reduced network quasi-steady-state species graph noninteracting dynamical system positivity

MSC Codes: MSC 92C42 MSC 80A30

1 Introduction

Biochemical reaction networks often involve many biochemical species that interact through many reactions. The mathematical models that are used to describe such networks can be quite complex and analytically intractable, both in terms of variables (species concentrations) as well as (unknown) parameters such as reaction rate constants. It is therefore commonplace to reduce the dimensions of the models by various means.

One way to simplify a model is by time-scale separation. Perhaps the first to formalise this idea were Briggs and Haldane in their study of the basic Michaelis-Menten enzyme-substrate mechanism [2], though Michaelis and Menten themselves had similar ideas [3]. Time-scale separation has found its use in many areas of theoretical biochemistry [10] as well as in modelling in general. To apply time-scale separation, the species in the system are divided into fast and slow species. The fast species are assumed to be at equilibrium (a so-called quasi-equilibrium) even though the system as a whole has not reached a state of equilibrium. For example, if the degradation of certain species proceed at high rate, then the species are short lived and become 'fast' species. It is thus reasonable to assume the species are in a state of (quasi-)equilibrium. Being at equilibrium, the fast species might be eliminated from the description of the system, resulting in a simpler system with only the slow species.

Our aim is to provide a *systematic* method to perform model reduction in a manner akin to reduction by time-scale separation, following some ideas introduced in [5, 6]. The method proceeds in two steps. The original system is described by a system of ordinary differential equations (ODEs). We abstract from the particular biological context and merely assume that a set of species (the fast species) is given and that these species effectively are at steady state.

That is, we do not make any biological assumptions or justifications to why this might be the case, hence there is no particular meaning attached to the term 'fast'. We further assume that the fast species are noninteracting in the sense that they do not form complexes with each other. If that is the case, the steady state equations for the fast species impose certain constraints on the relationship between the concentrations of the fast and the non-fast species [5]. Using this relationship, we derive an ODE system for the non-fast species in which the fast species have been eliminated.

The second step is an interpretation of the ODE system for the non-fast species as the ODE system associated with a *reduced reaction network*. This step is crucial and could be carried out in various ways. We propose a graphical method to determine the reduced reaction network, which can easily be carried out by hand for small networks. We will show that under reasonable assumptions, the kinetics of our reduced reaction network has desirable properties. For example, one might require that a reaction cannot occur in the absence of the reactant species. If the original reaction network fulfils this requirement, then so does our reduced reaction network.

Others have worked on similar ideas before us. The King-Altman procedure gives a systematic way to eliminate enzymes and intermediate species, in an enzyme catalysed system with mass-action kinetics [12]. Later, Wong and Hanes [20] gave a systematic method to find the production rate of the slow species, avoiding much of the algebraic manipulations in [12]. These approaches have been explained in a common framework [10]. A linear (graphical) elimination procedure generalising the above has also been proposed [5, 6].

None of the mentioned approaches specifically identify a reduced reaction network, except [6], where a reduced reaction network is identified after elimination of intermediate species. Another algebraic approach to reduce a reaction network by elimination of species has recently been proposed [13]. It is based on a graphical procedure to identify, not one but many, reduced reaction networks [15]. The procedure is similar in spirit to our method and we will contrast the two in Section 6.4.

We end this introduction by illustrating the aim of the paper in a particular example. Consider a *ping-pong bi-bi* mechanism [9], described by the reaction network

$$E + S_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} Y_1 \underset{k_4}{\overset{k_3}{\rightleftharpoons}} E^* + P_1 \qquad E^* + S_2 \underset{k_6}{\overset{k_5}{\rightleftharpoons}} Y_2 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} E + P_2,$$

where E, E^* are two forms of an enzyme, S_1, S_2 are two substrates, Y_1, Y_2 are two intermediate complexes and P_1, P_2 are two products. We aim to eliminate the set of species $\{E, E^*, Y_1, Y_2\}$. The molar concentrations of the species are denoted as

$$\begin{aligned} x_1 &= [E] & x_2 &= [E^*] & x_3 &= [S_1] & x_4 &= [S_2] \\ x_5 &= [P_1] & x_6 &= [P_2] & x_7 &= [Y_1] & x_8 &= [Y_2]. \end{aligned}$$

With this notation and assuming mass-action kinetics, the evolution of the species concentrations is described by the following ODE system:

$$\begin{aligned} \dot{x}_1 &= -(k_1 x_3 + k_8 x_6) x_1 + k_2 x_7 + k_7 x_8 & \dot{x}_2 &= -(k_4 x_5 + k_5 x_4) x_2 + k_3 x_7 + k_6 x_8 \\ \dot{x}_3 &= -k_1 x_1 x_3 + k_2 x_7 & \dot{x}_4 &= -k_5 x_2 x_4 + k_6 x_8 \\ \dot{x}_5 &= -k_4 x_2 x_5 + k_3 x_7 & \dot{x}_6 &= -k_8 x_1 x_6 + k_7 x_8 \\ \dot{x}_7 &= k_1 x_1 x_3 + k_4 x_2 x_5 - (k_2 + k_3) x_7 & \dot{x}_8 &= k_5 x_2 x_4 + k_8 x_1 x_6 - (k_6 + k_7) x_8. \end{aligned}$$

If we assume that E, E^*, Y_1, Y_2 are at steady state and use the conservation law for the total amount of enzyme, $T = x_1 + x_2 + x_7 + x_8$, then

$$\begin{aligned} x_1 =& q_T(x)((k_6 + k_7)k_2k_4x_5 + (k_2 + k_3)k_5k_7x_4) \\ x_2 =& q_T(x)((k_6 + k_7)k_1k_3x_3 + (k_2 + k_3)k_6k_8x_6) \\ x_7 =& q_T(x)(k_1k_5k_7x_3x_4 + (k_6 + k_7)k_1k_4x_3x_5 + k_4k_6k_8x_5x_6) \\ x_8 =& q_T(x)(k_1k_3k_5x_4x_3 + k_2k_4k_8x_5x_6 + (k_2 + k_3)k_5k_8x_4x_6), \end{aligned} \tag{1}$$

where $q_T(x)$ is a rational function depending on T and the non-eliminated variables.

After substitution of (1) into the ODE system, we obtain the production rates of the products P_1, P_2:

$$\dot{x}_5 = \dot{x}_6 = -\dot{x}_3 = -\dot{x}_4 = q_T(x)(-k_2k_4k_6k_8x_5x_6 + k_1k_3k_5k_7x_4x_3). \tag{2}$$

This ODE system might be interpreted as a reaction network with only one reversible reaction

$$S_1 + S_2 \rightleftharpoons P_1 + P_2.$$

The rate function for the forward reaction is $q_T(x)k_1k_3k_5k_7x_3x_4$ and for the backward reaction, $q_T(x)k_2k_4k_6k_8x_5x_6$.

In the particular case, it seems straight forward to identify reasonable reactions that explain (2). However, in general this might not be so. A particular problem is to make sure the kinetics is related to the reduced reactions in a sensible way, for example, it is natural to require that the reduced reaction rates retain 'mass-action form' in some sense, if the original kinetics were of mass-action type. In the example, the reduced reaction rates are mass-action apart from the denominator of $q_T(x)$, hence the reactions will only occur in the presence of reactants.

We will show how a reduced reaction network can be identified by graphical means. The graphical procedure further stipulates the reaction rates of the reduced reaction network.

The outline of the paper is the following. In Section 2 we introduce the notation and some background on graphs. In Section 3 we recall some earlier results on elimination of noninteracting species from [5]. In Section 4 and 5 we develop the reduction procedure and discuss properties of the reduced reaction networks, in particular in relation to the kinetics and conservation laws. Finally, in Section 6 we apply the procedure to different biological examples.

2 Preliminaries

In this section we introduce some concepts and notation from graph theory and chemical reaction network theory.

We let $\mathbb{R}_{\geq 0}$ and $\mathbb{R}_{>0}$ denote the sets of nonnegative reals numbers and positive real numbers respectively, and define $\mathbb{R}^n_{\geq 0}$ and $\mathbb{R}^n_{>0}$ accordingly. For $x, y \in \mathbb{R}^n$, $x \cdot y$ denotes the scalar product associated with the Euclidian norm. Further, $\langle v_1, \ldots, v_r \rangle$ denotes the vector subspace generated by $v_1, \ldots, v_r \in \mathbb{R}^n$.

2.1 Graphs

Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be a directed graph (digraph). By abuse of notation, we write $e \in \mathcal{G}$ whenever $e \in \mathcal{E}$. A **spanning tree** τ is a directed subgraph of $\mathcal{G}$ with node set $\mathcal{N}$ and such that the

underlying undirected graph is connected and acyclic. The spanning tree is *rooted* at the node N if N is the only node with no outgoing edges. We let $\Theta(N)$ be the set of spanning trees of $\mathcal{G}$ rooted at N.

The graph $\mathcal{G}$ is **strongly connected** if there is a directed path from N_1 to N_2 for any pair of nodes N_1, N_2. Any directed path from N_1 to N_2 in a strongly connected graph can be extended to a spanning tree rooted at N_2. A **cycle** is a closed directed path $N_{i_1} \to N_{i_2} \to \cdots \to N_{i_n} \to N_{i_1}$ with no repeated nodes apart from the initial and terminal nodes. We denote by $\Delta(\mathcal{G})$ the set of cycles in $\mathcal{G}$. By definition all cycles are directed.

If $\pi\colon \mathcal{E} \to R$ is a labeling of $\mathcal{G}$ with values in some ring R, then any subgraph $\mathcal{H}$ of $\mathcal{G}$ inherits a labeling from $\mathcal{G}$. We extend π to the set of subgraphs of $\mathcal{G}$ by defining

$$\pi(\mathcal{H}) = \prod_{e \in \mathcal{H}} \pi(e).$$

Definition 1. A **multidigraph** $\mathcal{G}$ is an ordered pair of finite sets $(\mathcal{N}, \mathcal{E})$ equipped with two functions:

$$s\colon \mathcal{E} \to \mathcal{N} \qquad t\colon \mathcal{E} \to \mathcal{N}.$$

The elements of $\mathcal{N}$ are called nodes, the elements of $\mathcal{E}$ are called edges, and the functions s, t are the source and target functions, respectively.

The function s assigns to each edge the source node of the edge and the function t assigns to each edge the target node of the edge. In a multidigraph both self-edges (edges e with $t(e) = s(e)$) and parallel edges (edges e_1, e_2 with $t(e_1) = t(e_2)$ and $s(e_1) = s(e_2)$) are possible.

Spanning trees, cycles and labels for a multidigraph are defined analogous to those of a digraph. Note that the unique spanning tree of a multidigraph consisting of one node and one self-edge is the node itself.

We associate a digraph $\widehat{\mathcal{G}}$ with a multidigraph $\mathcal{G}$ by removing self-edges and joining parallel edges:

Definition 2. Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be a labeled multidigraph and let $\mathcal{E}' = \{e \in \mathcal{E} \mid s(e) \neq t(e)\}$ be the set of edges that are not self-edges. The labeled digraph $\widehat{\mathcal{G}} = (\widehat{\mathcal{N}}, \widehat{\mathcal{E}})$ associated with $\mathcal{G}$ is the graph with

$$\widehat{\mathcal{N}} = \{N \in \mathcal{N} \mid \exists\, e \in \mathcal{E}' \text{ with } t(e) = N \text{ or } s(e) = N\}$$
$$\widehat{\mathcal{E}} = \{N_1 \to N_2 \mid \exists\, e \in \mathcal{E}' \text{ with } s(e) = N_1 \text{ and } t(e) = N_2\}.$$

We define a surjective map from $\mathcal{E}'$ to $\widehat{\mathcal{E}}$ as follows:

$$\begin{array}{rccc} \zeta\colon & \mathcal{E}' & \to & \widehat{\mathcal{E}} \\ & e & \mapsto & s(e) \to t(e). \end{array}$$

Using this map, a labeling $\widehat{\pi}$ for $\widehat{\mathcal{G}}$ can be defined as

$$\widehat{\pi}(\widehat{e}) = \sum_{e \in \zeta^{-1}(\widehat{e})} \pi(e).$$

Note that by definition $\widehat{\mathcal{N}} \subseteq \mathcal{N}$ and that the inclusion might be strict if there is a node in $\mathcal{G}$ that is only connected to itself.

Lemma 1. *Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be a labeled multidigraph and let $\widehat{\mathcal{G}} = (\widehat{\mathcal{N}}, \widehat{\mathcal{E}})$ be the associated digraph given in Definition 2. Let $N \in \widehat{\mathcal{N}}$ and let $\widehat{\Theta}(N)$, $\Theta(N)$ be the set of spanning trees rooted at N of $\widehat{\mathcal{G}}$ and $\mathcal{G}$, respectively. Then*

$$\sum_{\widehat{\tau} \in \widehat{\Theta}(N)} \widehat{\pi}(\widehat{\tau}) = \sum_{\tau \in \Theta(N)} \pi(\tau).$$

Proof. Since a spanning tree cannot contain a self-edge, the map ζ extends to a surjective map from $\Theta(N)$ to $\widehat{\Theta}(N)$. In particular,

$$\Theta(N) = \coprod_{\widehat{\tau} \in \widehat{\Theta}(N)} \zeta^{-1}(\widehat{\tau}). \tag{3}$$

Let $\widehat{\tau}$ be a spanning tree rooted at N in $\widehat{\mathcal{G}}$. Then,

$$\widehat{\pi}(\widehat{\tau}) = \prod_{\widehat{e} \in \widehat{\tau}} \widehat{\pi}(\widehat{e}) = \prod_{\widehat{e} \in \widehat{\tau}} \sum_{e \in \zeta^{-1}(\widehat{e})} \pi(e) = \sum_{\tau \in \zeta^{-1}(\widehat{\tau})} \prod_{e \in \tau} \pi(e) = \sum_{\tau \in \zeta^{-1}(\widehat{\tau})} \pi(\tau).$$

Therefore, using (3) we obtain

$$\sum_{\tau \in \Theta(N)} \pi(\tau) = \sum_{\widehat{\tau} \in \widehat{\Theta}(N)} \sum_{\tau \in \zeta^{-1}(\widehat{\tau})} \pi(\tau) = \sum_{\widehat{\tau} \in \widehat{\Theta}(N)} \widehat{\pi}(\widehat{\tau}).$$

□

2.2 Reaction networks

Definition 3. A **reaction network** on a finite set $\mathcal{S}$ is a multidigraph $(\mathcal{C}, \mathcal{R})$ where

(i) $\mathcal{S} = \{S_1, \ldots, S_n\}$ is called the **species** set.

(ii) $\mathcal{C} \subset \mathbb{R}^n_{\geq 0}$ is called the set of **complexes**.

(iii) $\mathcal{R} = \{r_1, \ldots, r_\ell\}$ is called the set of **reactions**.

The source and target functions (Definition 1) are adapted to the nomenclature of reaction networks as follows:

- The source function is denoted as $y \colon \mathcal{R} \to \mathcal{C}$ and assigns to each reaction r_i its **reactant** y_{r_i}.
- The target function is denoted as $y' \colon \mathcal{R} \to \mathcal{C}$ and assigns to each reaction r_i its **product** y'_{r_i}.

Further, we assume that $y_r \neq y'_r$ for all reactions $r \in \mathcal{R}$, and that every complex $\eta \in \mathcal{C}$ is either the source or target of some reaction $r \in \mathcal{R}$. That is, the multidigraph $(\mathcal{C}, \mathcal{R})$ contains no self-edges.

Our definition of reaction network differs from the usual one in that we allow for multiple reactions with the same reactant and product (e.g. [4]). We do this for mathematical convenience as will be clear in Section 4. Note, however, that if $\mathcal{G}$ is a reaction network, then $\widehat{\mathcal{G}}$ (Definition 2) is a reaction network in the standard sense.

It is custom to express a complex $\eta \in \mathcal{C}$ as a linear combination of the species, in particular, this is the case in the examples. For example, the complex $(1, 0, 1) \in \mathcal{C}$ is written as $S_1 + S_3$.

Definition 4. Let $\eta \in \mathcal{C} \subset \mathbb{R}^n_{\geq 0}$ be a complex of a reaction network $(\mathcal{C}, \mathcal{R})$. We use the following terminology:

(1) The ith component, η_i, of η is the **stoichiometric coefficient** of S_i in η.

(2) If $\eta_i \neq 0$ for some i, then η **involves** S_i and r **involves** S_i for any reaction r such that $y_r = \eta$ or $y'_r = \eta$.

(3) A pair of species $S_i, S_j \in \mathcal{S}$, $i \neq j$, **interact** if $\eta_i, \eta_j \neq 0$, for some complex η. Equivalently, if they are both involved in the same reactant or product of a reaction.

We further define:

(4) The **stoichiometric matrix** is the matrix $A = (a_{ij}) \in \mathbb{R}^{n \times \ell}$ where $a_{ij} = (y'_{r_j} - y_{r_j})_i$.

(5) The **support** of a vector $\omega = (\omega_1, \ldots, \omega_n) \in \mathbb{R}^n$ is the set

$$\operatorname{supp}(\omega) = \{S_i \mid \omega_i \neq 0\} \subseteq \mathcal{S}.$$

Although a reaction network is defined as a multidigraph, we often give a reaction network by listing the set of reactions $\mathcal{R}$. The set of complexes $\mathcal{C}$ and the set of species $\mathcal{S}$ can then easily be extracted.

Example 5. The following

$$S_1 + S_4 \rightleftharpoons S_5 \longrightarrow S_2 + S_4 \longrightarrow S_3 + S_4$$

is a reaction network on the set of species $\mathcal{S} = \{S_1, S_2, S_3, S_4, S_5\}$ with set of complexes $\mathcal{C} = \{S_1 + S_4, S_5, S_2 + S_4, S_3 + S_4\}$.

We denote by $x_i = x_i(t)$ the concentration of species S_i at time t and let $x = (x_1, \ldots, x_n)$ be the vector of concentrations. In specific examples, where the species are denoted by letters such as E, P, Q, the concentrations are denoted by x_E, x_P, x_Q, respectively.

Definition 6. A **kinetics** for a reaction network $\mathcal{G} = (\mathcal{C}, \mathcal{R})$ is a function

$$\begin{array}{rccl} \kappa\colon & \Omega & \to & \mathbb{R}^\ell_{\geq 0} \\ & x & \mapsto & (\kappa_{r_1}(x), \ldots, \kappa_{r_\ell}(x)), \end{array}$$

where $\mathbb{R}^n_{>0} \subseteq \Omega \subseteq \mathbb{R}^n_{\geq 0}$, such that $\kappa(\mathbb{R}^n_{>0}) \subseteq \mathbb{R}^\ell_{>0}$. The component $\kappa_{r_i}(x)$ of $\kappa(x)$ is called the **rate function** of reaction $r_i \in \mathcal{R}$.

Mass-action kinetics is the kinetics defined by the following rate functions:

$$\kappa_r(x) = k_r x^{y_r} = k_r \prod_{i=1}^{n} x_i^{(y_r)_i}, \qquad r \in \mathcal{R},$$

where $k_r > 0$ is a constant called **the reaction rate constant** of reaction r. By convention, $0^0 = 1$.

Remark 7. The kinetics κ provides a labeling of the reaction network. Since the set of reactions is ordered, we often denote the rate functions as $\kappa_i(x)$ instead of $\kappa_{r_i}(x)$. In the examples, the labeling implicitly indicates the chosen order. In the particular case of mass-action kinetics, we simply use the reaction rate constants k_i as labels.

Given a reaction network $\mathcal{G} = (\mathcal{C}, \mathcal{R})$ on $\mathcal{S}$ and a kinetics $(\kappa_{r_1}(x), \ldots, \kappa_{r_\ell}(x))$, we let

$$g_i(x) = \sum_{r \in \mathcal{R}} \kappa_r(x)(y'_r - y_r)_i, \qquad i = 1, \ldots, n, \quad x \in \Omega. \tag{4}$$

The evolution of the species concentration in time is modelled by the following system of ODEs:

$$\dot{x}_i = g_i(x), \qquad i = 1, \ldots, n, \qquad x \in \Omega. \tag{5}$$

By letting $\dot{x} = (\dot{x}_1, \ldots, \dot{x}_n)$ and $g(x) = (g_1(x), \ldots, g_n(x))$, the ODE system can be written in vector notation as $\dot{x} = g(x)$, $x \in \Omega$, or using the specific form of the equations (4), as

$$g(x) = A\kappa(x), \quad x \in \Omega. \tag{6}$$

The **steady states** of the system (5) are the solutions to the system

$$g(x) = 0, \qquad x \in \Omega.$$

Remark 8. Consider a reaction network $\mathcal{G}$ with a kinetics κ and the associated reaction network $\widehat{\mathcal{G}}$ with the induced kinetics $\widehat{\kappa}$ (cf. the induced labeling in Definition 2). Then the ODE system associated with $(\mathcal{G}, \kappa)$ agrees with the ODE system associated with $(\widehat{\mathcal{G}}, \widehat{\kappa})$.

Remark 9. In typical models of biochemical reaction systems, the rate function of a reaction vanishes whenever the concentration of one of the reactant species is zero. In particular, it guarantees invariance of the non-negative orthant under (5). However, we do not need this assumption for our results to hold. In Section 4.4 we discuss some results that follow from making this assumption.

The **stoichiometric subspace** of a network $(\mathcal{C}, \mathcal{R})$ is the following vector subspace of $\mathbb{R}^n$:

$$S = \langle y'_r - y_r \mid r \in \mathcal{R} \rangle \subset \mathbb{R}^n.$$

If $\omega = (\omega_1, \ldots, \omega_n) \in S^\perp$, then it follows from (5) and (4) that $\omega \cdot \dot{x} = 0$. Thus $\omega \cdot x$ is independent of time along any trajectory and we say that

$$\omega \cdot x = \sum_{i=1}^{n} \omega_i x_i$$

is **conserved**. In particular, for any trajectory, there is a constant $T \in \mathbb{R}$ such that

$$T = \omega \cdot x.$$

This equation is called the **conservation law** with total amount $T \in \mathbb{R}$, corresponding to $\omega \in S^\perp$.

We say that a set of conservation laws is **minimal** if it is derived from a basis $\omega^1, \ldots, \omega^l$ of $S^\perp$. Then, the trajectory with initial concentrations x_0 is confined to the linear space with equations

$$T_1 = \omega^1 \cdot x, \ \ldots \ , \ T_l = \omega^l \cdot x,$$

where $T_i = \omega^i \cdot x_0$, for $i = 1, \ldots, l$. The intersection of this linear space with $\mathbb{R}^n_{\geq 0}$ is called a **stoichiometric compatibility class** of the system.

Example 10. Consider Example 5 and let κ be a kinetics:

$$S_1 + S_4 \underset{\kappa_2(x)}{\overset{\kappa_1(x)}{\rightleftharpoons}} S_5 \xrightarrow{\kappa_3(x)} S_2 + S_4 \xrightarrow{\kappa_4(x)} S_3 + S_4.$$

The corresponding ODE system is:

$$\begin{aligned} \dot{x}_1 &= -\kappa_1(x) + \kappa_2(x), & \dot{x}_2 &= \kappa_3(x) - \kappa_4(x), & \dot{x}_3 &= \kappa_4(x), \\ \dot{x}_4 &= -\kappa_1(x) + \kappa_2(x) + \kappa_3(x), & \dot{x}_5 &= \kappa_1(x) - \kappa_2(x) - \kappa_3(x), \end{aligned}$$

and a minimal set of conservation laws consists of

$$x_1 + x_2 + x_3 + x_5 = T_1, \qquad x_4 + x_5 = T_2. \tag{7}$$

The stoichiometric compatibility class is non-empty, if and only if $T_1, T_2 \in \mathbb{R}_{\geq 0}$.

3 Elimination of variables

In this section we recall some results on linear elimination of variables from [5]. We assume a reaction network $(\mathcal{C}, \mathcal{R})$ on $\mathcal{S}$ is given.

Definition 11. A subset $\mathcal{U} \subset \mathcal{S}$ is **noninteracting** if it contains no pair of interacting species, and the stoichiometric coefficients of the species in $\mathcal{U}$ in all complexes are either 0 or 1.

Let $\mathcal{U} \subset \mathcal{S}$ be a noninteracting subset of species. For simplicity we let $\mathcal{U} = \{U_1, \ldots, U_m\}$ and $\mathcal{U}^c = \mathcal{S} \setminus \mathcal{U} = \{S_1, \ldots, S_p\}$ (with $p = n - m$), such that $\mathcal{S} = \mathcal{U}^c \cup \mathcal{U}$. We order $\mathcal{S}$ as

$$\mathcal{S} = \{S_1, \ldots, S_p, U_1, \ldots, U_m\}.$$

From now on, we let x_i be the concentration of $S_i \in \mathcal{U}^c$ and $x = (x_1, \ldots, x_p)$. Similarly, we let u_i be the concentration of $U_i \in \mathcal{U}$ and $u = (u_1, \ldots, u_m)$. Hence, the rate functions are functions of (x, u): $\kappa_r(x, u)$.

Definition 12. A noninteracting set $\mathcal{U} \subset \mathcal{S}$ is a **cut** if $\omega = (0, \overset{p}{\ldots}, 0, 1, \overset{m}{\ldots}, 1) \in S^{\perp}$. That is, if $(1, \overset{m}{\ldots}, 1) \cdot u = \sum_{i=1}^{m} u_i$ is conserved.

The set $\mathcal{U} = \{S_4, S_5\}$ in Example 5 is noninteracting and a cut, see (7).

Let $\mathcal{U} \subset \mathcal{S}$ be a noninteracting set and let

$$\rho_c \colon \mathbb{R}^n \to \mathbb{R}^p, \qquad \rho \colon \mathbb{R}^n \to \mathbb{R}^m$$

be the projection onto the first p coordinates and last m coordinates of $\mathbb{R}^n$, respectively. We define the following subsets of $\mathcal{R}$:

(i) $\mathcal{R}_{\mathcal{U}}$ is the set of reactions that involve species in $\mathcal{U}$ in the reactant as well as in the product:

$$\mathcal{R}_{\mathcal{U}} = \{r \in \mathcal{R} \mid \rho(y_r) \neq 0 \text{ and } \rho(y_r') \neq 0\}.$$

(ii) $\mathcal{R}^c_{\mathcal{U},y}$ is the set of reactions that involve species in $\mathcal{U}$ in the reactant but not in the product:

$$\mathcal{R}^c_{\mathcal{U},y} = \{r \in \mathcal{R} \mid \rho(y_r) \neq 0 \text{ and } \rho(y_r') = 0\}.$$

(iii) $\mathcal{R}^c_{\mathcal{U},y'}$ is the set of reactions that involve species in $\mathcal{U}$ in the product but not in the reactant:

$$\mathcal{R}^c_{\mathcal{U},y'} = \{r \in \mathcal{R} \mid \rho(y'_r) \neq 0 \text{ and } \rho(y_r) = 0\}.$$

(iv) $\mathcal{R}^\perp_{\mathcal{U}}$ is the set of reactions that do not involve any species in $\mathcal{U}$:

$$\mathcal{R}^\perp_{\mathcal{U}} = \{r \in \mathcal{R} \mid \rho(y_r) = \rho(y'_r) = 0\}.$$

We further define

$$\mathcal{R}^c_{\mathcal{U}} = \mathcal{R}^c_{\mathcal{U},y} \cup \mathcal{R}^c_{\mathcal{U},y'}.$$

Clearly,

$$\mathcal{R} = \mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^\perp_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U}}.$$

Moreover, $\mathcal{R}^c_{\mathcal{U}} = \emptyset$ if and only if $\mathcal{U}$ is a cut. Indeed, since $(1, \dots, 1) \cdot u$ is conserved, any reaction that involves a species in $\mathcal{U}$ in the reactant must also involve one in the product.

If $\mathcal{U}$ is noninteracting, the vector $\rho(y_r)$, respectively $\rho(y'_r)$, has exactly one nonzero component if $r \in \mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U},y}$, respectively, if $r \in \mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U},y'}$.

Definition 13. Let $\kappa\colon \Omega \times \mathbb{R}^m_{\geq 0} \to \mathbb{R}^\ell_{\geq 0}$ be a kinetics with $\mathbb{R}^p_{>0} \subseteq \Omega \subseteq \mathbb{R}^p_{\geq 0}$. The kinetics κ is **$\mathcal{U}$-linear** if, for each $r \in \mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U}}$, there exists a function $v_r\colon \Omega \to \mathbb{R}_{\geq 0}$ such that $v_r(x) > 0$ for all $x \in \mathbb{R}^p_{>0}$ and

$$\kappa_r(x,u) = \begin{cases} u_i v_r(x) & \text{if } r \in \mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U},y} \text{ and } \rho(y_r)_i \neq 0 \\ v_r(x) & \text{if } r \in \mathcal{R}^c_{\mathcal{U},y'}. \end{cases}$$

We do not make any additional assumptions on the rate functions for $r \in \mathcal{R}^\perp_{\mathcal{U}}$.

Definition 14. Let $\mathcal{U} \subset \mathcal{S}$ be a set of noninteracting species and κ a $\mathcal{U}$-linear kinetics. We define the labeled multidigraph $\mathcal{G}_{\mathcal{U}} = (\mathcal{N}_{\mathcal{U}}, \mathcal{E}_{\mathcal{U}})$ as follows:

- If $\mathcal{U}$ is a cut, $\mathcal{N}_{\mathcal{U}} = \mathcal{U}$ and

$$\mathcal{E}_{\mathcal{U}} = \left\{U_i \xrightarrow{v_r(x)} U_j \mid r \in \mathcal{R}_{\mathcal{U}} \text{ with } \rho(y_r)_i \neq 0 \text{ and } \rho(y'_r)_j \neq 0\right\}.$$

- If $\mathcal{U}$ is not a cut, $\mathcal{N}_{\mathcal{U}} = \mathcal{U} \cup \{*\}$ and

$$\begin{aligned} \mathcal{E}_{\mathcal{U}} = &\{U_i \xrightarrow{v_r(x)} U_j \mid r \in \mathcal{R}_{\mathcal{U}} \text{ with } \rho(y_r)_i \neq 0 \text{ and } \rho(y'_r)_j \neq 0\} \cup \\ &\{U_i \xrightarrow{v_r(x)} * \mid r \in \mathcal{R}^c_{\mathcal{U},y} \text{ with } \rho(y_r)_i \neq 0\} \cup \\ &\{* \xrightarrow{v_r(x)} U_i \mid r \in \mathcal{R}^c_{\mathcal{U},y'} \text{ with } \rho(y'_r)_i \neq 0\}. \end{aligned}$$

Each edge in the multidigraph $\mathcal{G}_{\mathcal{U}}$ corresponds to a reaction in $\mathcal{R}$. Moreover, since $\mathcal{U}$ is a noninteracting set, the multidigraph $\mathcal{G}_{\mathcal{U}}$ has exactly one edge for each reaction in $\mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U}}$ and no edges corresponding to reactions in $\mathcal{R}^\perp_{\mathcal{U}}$. This gives rise to two bijective functions

$$\mathcal{E}_{\mathcal{U}} \underset{e}{\overset{r}{\rightleftarrows}} \mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U}}, \tag{8}$$

such that $r \circ e = \mathrm{id}_{\mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U}}}$ and $e \circ r = \mathrm{id}_{\mathcal{E}_{\mathcal{U}}}$. The function r maps an edge of the multidigraph $\mathcal{G}_{\mathcal{U}}$ to the corresponding reaction and the function e maps a reaction in $\mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U}}$ to the

corresponding edge. In the examples, the functions r, e are implicitly given by the subindices of the functions $v_r(x)$: the edge label $v_i(x)$ indicates that the edge corresponds to the reaction r_i.

Observe that $\mathcal{G}_{\mathcal{U}}$ might contain parallel edges between any pair of nodes and self-edges for nodes different from $*$.

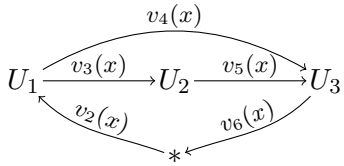
Example 15. Consider the network

$$S_1 \xrightarrow{\kappa_1(x,u)} S_2 \xrightarrow{\kappa_2(x,u)} U_1 \xrightarrow{\kappa_3(x,u)} U_2 \qquad S_3 + U_1 \xrightarrow{\kappa_4(x,u)} U_3 + S_1$$
$$S_3 + U_2 \xrightarrow{\kappa_5(x,u)} S_2 + U_3 \qquad U_3 \xrightarrow{\kappa_6(x,u)} S_3$$

and $\mathcal{U} = \{U_1, U_2, U_3\}$. Then $\mathcal{U}$ is a noninteracting set that is not a cut. A $\mathcal{U}$-linear kinetics takes the form

$$\kappa_2(x,u) = v_2(x), \qquad \kappa_3(x,u) = u_1 v_3(x), \qquad \kappa_4(x,u) = u_1 v_4(x),$$
$$\kappa_5(x,u) = u_2 v_5(x), \qquad \kappa_6(x,u) = u_3 v_6(x),$$

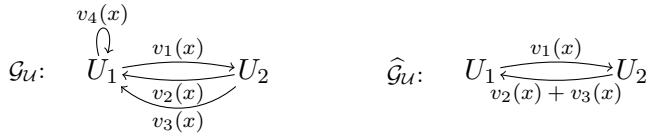
and the corresponding multidigraph $\mathcal{G}_{\mathcal{U}}$ is



Example 16. Consider the network in Example 5 and the cut $\mathcal{U} = \{S_4, S_5\}$. To adapt the notation to the present setting, we let $U_1 = S_4$ and $U_2 = S_5$. A $\mathcal{U}$-linear kinetics satisfies

$$\kappa_1(x,u) = u_1 v_1(x), \qquad \kappa_2(x,u) = u_2 v_2(x),$$
$$\kappa_3(x,u) = u_2 v_3(x), \qquad \kappa_4(x,u) = u_1 v_4(x).$$

The multidigraph $\mathcal{G}_{\mathcal{U}}$ and the digraph $\widehat{\mathcal{G}}_{\mathcal{U}}$ are given by



The graph $\mathcal{G}_{\mathcal{U}}$ is connected if and only if $\widehat{\mathcal{G}}_{\mathcal{U}}$ is connected. Therefore, we might reformulate the following result from [5] in terms of $\mathcal{G}_{\mathcal{U}}$.

Lemma 2. *Assume that $\mathcal{G}_{\mathcal{U}}$ is connected.*

(i) If $\mathcal{U}$ is a cut, then any vector in $S^{\perp}$ with support included in $\mathcal{U}$ is a multiple of $\omega = (0, \overset{p}{\ldots}, 0, 1, \overset{m}{\ldots}, 1)$.

(ii) If $\mathcal{U}$ is not a cut, then $\mathcal{U}$ does not include the support of any vector in $S^{\perp}$.

Henceforth, we will assume that $\mathcal{G}_{\mathcal{U}}$ is connected, unless otherwise specified. Consider the following assumption:

(ST) There exists a spanning tree of $\mathcal{G}_{\mathcal{U}}$ rooted at some $U_i \in \mathcal{U}$, if $\mathcal{U}$ is a cut, and rooted at $*$, if $\mathcal{U}$ is not a cut.

Let $\Theta(N)$ denote the set of spanning trees in $\mathcal{G}_\mathcal{U}$ rooted at $N \in \mathcal{N}_\mathcal{U}$. Let a stoichiometric compatibility class be given such that, if $\mathcal{U}$ is a cut, then

$$T = \sum_{i=1}^{m} u_i, \qquad T \geq 0.$$

If (ST) holds, we define the function $q(x)$ by:

$$q(x) = \frac{T}{\sum\limits_{i=1}^{m} \sum\limits_{\tau \in \Theta(U_i)} \pi(\tau)}, \qquad \text{if } \mathcal{U} \text{ is a cut,} \tag{9}$$

$$q(x) = \frac{1}{\sum\limits_{\tau \in \Theta(*)} \pi(\tau)}, \qquad \text{if } \mathcal{U} \text{ is not a cut.} \tag{10}$$

The function $q(x)$ is a function of x since $\pi(\tau)$ depends on x. We omit the explicit reference to $\mathcal{U}$ in $q(x)$ and to x in $\pi(\tau)$ for convenience. If the spanning tree τ contains no edges (that is, $\mathcal{G}_\mathcal{U}$ consists of exactly one node and one self-edge), then by definition $\pi(\tau) = 1$.

Since the functions $v_r(x)$ are positive in $\mathbb{R}^p_{>0}$, the function $q(x)$ is positive for all $x \in \mathbb{R}^p_{>0}$. However, the denominator of $q(x)$ is defined for all $x \in \Omega$ and might vanish at points at the boundary of Ω. In that case $q(x)$ is not well defined. Henceforth, we let $\mathbb{R}^p_{>0} \subseteq \widetilde{\Omega} \subseteq \Omega$ be the domain where $q(x)$ is well defined.

Using Lemma 1, we might reformulate Propositions 8.4 and 8.6 in [5] in terms of the multidigraph $\mathcal{G}_\mathcal{U}$ for a reaction network with $\mathcal{U}$-linear kinetics.

Theorem 17 (Elimination Theorem [5])**.** *Let $\mathcal{U} \subset \mathcal{S}$ be a set of noninteracting species and κ a $\mathcal{U}$-linear kinetics. Assume that $\mathcal{G}_\mathcal{U}$ is connected and that (ST) holds. Then, the system of equations*

$$\dot{u} = 0, \quad \textit{and} \quad T = (1, \ldots, 1) \cdot u, \quad T \geq 0$$

(the latter equation, only if $\mathcal{U}$ is a cut) holds for $(x, u) \in \widetilde{\Omega} \times \mathbb{R}^m_{\geq 0}$ if and only if

$$u_i = q(x) \sum_{\tau \in \Theta(U_i)} \pi(\tau), \qquad \textit{for all } i = 1 \ldots m.$$

For future reference, we define the function $\varphi \colon \widetilde{\Omega} \to \mathbb{R}^m_{\geq 0}$:

$$\varphi(x) = (\varphi_1(x), \ldots, \varphi_m(x)), \qquad \varphi_i(x) = q(x) \sum_{\tau \in \Theta(U_i)} \pi(\tau). \tag{11}$$

The theorem states that under the assumption that the species in $\mathcal{U}$ are at steady state with respect to the species in $\mathcal{U}^c$, the concentrations u_i are expressed as rational functions in the labels of $\mathcal{G}_\mathcal{U}$. The labels depend on the concentrations of the species in $\mathcal{U}^c$ only. In virtue of the non-negativity of the labels of $\mathcal{G}_\mathcal{U}$, the concentrations of the species in $\mathcal{U}$ are also non-negative for $x \in \widetilde{\Omega}$. We say that the species in $\mathcal{U}$ can be **eliminated**.

If the graph $\mathcal{G}_\mathcal{U}$ is strongly connected, then there is at least one spanning tree rooted at each node, and hence $\varphi(x)$ has no components that are identically zero. If the graph $\mathcal{G}_\mathcal{U}$ is not strongly connected, then at least one of the nodes is not the root of any spanning tree, and hence at least one of the components of $\varphi(x)$ is identically zero. In particular, in the later case, all steady states lay in the boundary of the positive orthant.

Example 18. We continue with Example 16 and eliminate the cut $\mathcal{U} = \{U_1, U_2\}$. By Theorem 17, we obtain

$$u_1 = \frac{T_2\,(v_3(x) + v_2(x))}{v_1(x) + v_2(x) + v_3(x)}, \qquad u_2 = \frac{T_2 v_1(x)}{v_1(x) + v_2(x) + v_3(x)}.$$

Example 19. Consider now Example 15 with $\mathcal{U} = \{U_1, U_2, U_3\}$, which is not a cut. By Theorem 17, we find

$$\begin{aligned} u_1 &= q(x)v_2(x)v_5(x)v_6(x), & u_2 &= q(x)v_2(x)v_3(x)v_6(x),\\ u_3 &= q(x)v_2(x)(v_3(x) + v_4(x))v_5(x), & q(x) &= \frac{1}{(v_3(x) + v_4(x))v_5(x)v_6(x)}. \end{aligned}$$

4 Reduced reaction network

In this section we assume that a reaction network $(\mathcal{C}, \mathcal{R})$, a set of noninteracting species $\mathcal{U} \subset \mathcal{S}$ and a $\mathcal{U}$-linear kinetics κ are given. Further, we assume that the multidigraph $\mathcal{G}_{\mathcal{U}}$ is connected and fulfils assumption (ST).

If we substitute u by $\varphi(x)$ in the ODE system (5), then $\dot{u} = 0$ by construction and

$$\dot{x} = \tilde{g}(x), \qquad x \in \widetilde{\Omega} \subseteq \mathbb{R}^p_{\geq 0}, \tag{12}$$

where

$$\tilde{g}(x) = \rho_c(g(x, \varphi(x))).$$

That is, for $i = 1, \ldots, p$

$$\tilde{g}_i(x) = g_i\left(x_1, \ldots, x_p, q(x)\sum_{\tau \in \Theta(U_1)} \pi(\tau), \ldots, q(x)\sum_{\tau \in \Theta(U_m)} \pi(\tau)\right).$$

In this section we prove that (12) is the ODE system of a reaction network, by graphically defining the reactions of the network and their rate functions. The multidigraph $\mathcal{G}_{\mathcal{U}}$ is one of the main tools for doing so.

4.1 Graphical tools

Using the notation introduced in Section 2, we define the following set of sub-multidigraphs of $\mathcal{G}_{\mathcal{U}}$:

$$\Gamma := \big\{\gamma = \tau \cup e \;\mid\; \tau \in \Theta(s(e)), \quad e \in \mathcal{G}_{\mathcal{U}}\big\}. \tag{13}$$

Each element of Γ is the union of a spanning tree rooted at a node $N \in \mathcal{N}_{\mathcal{U}}$ and an edge $e \in \mathcal{E}_{\mathcal{U}}$ with source $s(e) = N$. Any multidigraph $\gamma \in \Gamma$ contains a unique cycle. Indeed, for $\gamma = \tau \cup e$ as in (13), the cycle is obtained by joining e and the path in τ from $t(e)$ to $s(e)$, which exists because τ is rooted at $s(e)$. In particular, the cycle contains the edge e.

Let $\Delta := \Delta(\mathcal{G}_{\mathcal{U}})$ be the set of cycles in the multidigraph $\mathcal{G}_{\mathcal{U}}$. We define $\Gamma(\sigma)$ as the subset of Γ whose elements contain $\sigma \in \Delta$:

$$\Gamma(\sigma) = \{\gamma \in \Gamma \mid \sigma \subset \gamma\}.$$

Note that $\Gamma = \coprod_{\sigma \in \Delta} \Gamma(\sigma)$ and that the union is disjoint. The following proposition simplifies the computation of $\Gamma(\sigma)$.

Proposition 20. *Let $\sigma \in \Delta$ be a cycle of $\mathcal{G}_{\mathcal{U}}$. For any edge $e \in \sigma$,*

$$\Gamma(\sigma) = \{\tau \cup e \mid \sigma \setminus e \subset \tau,\ \tau \in \Theta(s(e))\}.$$

In particular, if $\mathcal{G}_{\mathcal{U}}$ is strongly connected, then $\Gamma(\sigma) \neq \emptyset$.

Proof. If $\gamma \in \Gamma(\sigma)$, then $\gamma = \tau \cup e$ with $\tau \in \Theta(s(e))$ and $e \in \sigma$. Clearly τ is a spanning tree that contains $\sigma \setminus e$. Conversely, if $\tau \in \Theta(s(e))$ is a spanning tree containing $\sigma \setminus e$, then $\tau \cup e \in \Gamma(\sigma)$.

If $\mathcal{G}_{\mathcal{U}}$ is strongly connected, then $\Gamma(\sigma) \neq \emptyset$ because for any path $\sigma \setminus e$, there exists a spanning tree rooted at the terminal node $s(e)$. □

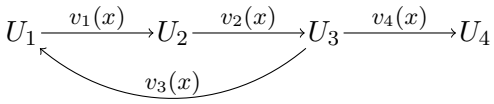
It follows from the proposition that in order to compute $\Gamma(\sigma)$, it is sufficient to consider any edge $e \in \sigma$ and find all spanning trees rooted at $s(e)$ containing $\sigma \setminus e$. We have that $\Gamma(\sigma) \neq \emptyset$ if and only if the path $\sigma \setminus e$ can be extended to a spanning tree rooted at $s(e)$, for any edge e of σ.

Remark 21. If σ is a cycle that contains all nodes of $\mathcal{G}_{\mathcal{U}}$, then $\Gamma(\sigma) = \{\sigma\}$. Indeed, in this case $\sigma \setminus e$ is the unique spanning tree of $\mathcal{G}_{\mathcal{U}}$ containing $\sigma \setminus e$.

Example 22. Consider the following reaction network:

$$\begin{aligned} S_1 + U_1 &\xrightarrow{\kappa_1(x,u)} U_2 \xrightarrow{\kappa_2(x,u)} S_2 + U_3 \\ S_3 + U_3 &\xrightarrow{\kappa_3(x,u)} S_4 + U_1 \\ S_4 + U_3 &\xrightarrow{\kappa_4(x,u)} S_3 + U_4 \end{aligned}$$

with the cut $\mathcal{U} = \{U_1, U_2, U_3, U_4\}$ and a $\mathcal{U}$-linear kinetics. The multidigraph $\mathcal{G}_{\mathcal{U}}$ is



for which (ST) holds. Consider the unique cycle σ of $\mathcal{G}_{\mathcal{U}}$. The graph $\sigma \setminus e$ with $e\colon U_1 \to U_2$ cannot be extended to a spanning tree rooted at $s(e) = U_1$. Therefore, $\Gamma(\sigma) = \emptyset$ by Proposition 20.

For $\sigma \in \Delta$, we define the function $\Pi(\sigma) := \sum_{\gamma \in \Gamma(\sigma)} \pi(\gamma)$. Note that

$$\Pi(\sigma) = \pi(\sigma) \sum_{\gamma \in \Gamma(\sigma)} \pi(\gamma \setminus \sigma). \tag{14}$$

4.2 Reduced reaction network

Let $\widetilde{\Delta}$ be the subset of Δ consisting of the cycles σ such that

$$\sum_{e \in \sigma} \rho_c(y'_{r(e)} - y_{r(e)}) \neq 0, \quad \text{and} \quad \Gamma(\sigma) \neq \emptyset.$$

The first condition states that the net-production of some species in $\mathcal{U}^c$ is non-zero over the reactions associated with the edges of the cycle.

Definition 23. Let $(\mathcal{C},\mathcal{R})$ be a reaction network on a set of species $\mathcal{S}$, $\mathcal{U}\subseteq\mathcal{S}$ a set of noninteracting species, and κ a $\mathcal{U}$-linear kinetics defined on $\Omega\times\mathbb{R}^m_{\geq 0}$, such that the multidigraph $\mathcal{G}_{\mathcal{U}}$ is connected and fulfils assumption (ST). The **reduced reaction network obtained by elimination of** $\mathcal{U}$ is the reaction network $(\widetilde{\mathcal{C}},\widetilde{\mathcal{R}})$ on the set of species $\mathcal{U}^c$ with kinetics $\widetilde{\kappa}$ defined on $\widetilde{\Omega}$, such that $\widetilde{\mathcal{R}}=\widetilde{\mathcal{R}}_1\cup\widetilde{\mathcal{R}}_2$, where

$$\widetilde{\mathcal{R}}_1=\left\{\widetilde{r}\colon \rho_c(y_r)\to\rho_c(y'_r)\ \mid\ r\in\mathcal{R}^{\perp}_{\mathcal{U}}\right\}$$
$$\widetilde{\mathcal{R}}_2=\left\{\widetilde{r}_\sigma\colon \sum_{e\in\sigma}\rho_c(y_{r(e)})\to\sum_{e\in\sigma}\rho_c(y'_{r(e)})\ \mid\ \sigma\in\widetilde{\Delta}\right\},$$

$\widetilde{\mathcal{C}}$ is the set of source and target nodes of $\widetilde{\mathcal{R}}$,

$$\widetilde{\kappa}_{\widetilde{r}}(x):=\begin{cases}\kappa_r(x,\varphi(x)) & \text{if } \widetilde{r}=\rho_c(y_r)\to\rho_c(y'_r)\in\widetilde{\mathcal{R}}_1,\ \text{for } r\in\mathcal{R}^{\perp}_{\mathcal{U}},\\ q(x)\Pi(\sigma) & \text{if } \widetilde{r}=\widetilde{r}_\sigma\in\widetilde{\mathcal{R}}_2.\end{cases}$$

and $\widetilde{\Omega}\subseteq\Omega$ is the set for which $\widetilde{\kappa}$ is well defined.

We are now ready to present the main result of the paper: the interpretation of the ODE system (12) as the ODE system of the reduced reaction network obtained by elimination of $\mathcal{U}$ defined above.

Theorem 24 (**Reduced reaction network).** *Let $(\mathcal{C},\mathcal{R})$ be a reaction network on a set species $\mathcal{S}$, $\mathcal{U}\subseteq\mathcal{S}$ a set of noninteracting species, and κ a $\mathcal{U}$-linear kinetics defined on $\Omega\times\mathbb{R}^m_{\geq 0}$, such that the multidigraph $\mathcal{G}_{\mathcal{U}}$ is connected and fulfils assumption (ST).*

Then the ODE system (12) *is the ODE system associated with the reduced reaction network obtained by elimination of $\mathcal{U}$.*

Proof. Let

$$f(x)=\sum_{\widetilde{r}\in\widetilde{\mathcal{R}}_1}\widetilde{\kappa}_{\widetilde{r}}(x)(y'_{\widetilde{r}}-y_{\widetilde{r}})+\sum_{\widetilde{r}\in\widetilde{\mathcal{R}}_2}\widetilde{\kappa}_{\widetilde{r}}(x)(y'_{\widetilde{r}}-y_{\widetilde{r}}).$$

We want to prove that $f(x)=\widetilde{g}(x)=\rho_c(g(x,\varphi(x)))$. Observe that

$$\widetilde{g}(x)=\sum_{r\in\mathcal{R}^{\perp}_{\mathcal{U}}}\kappa_r(x,\varphi(x))\rho_c(y'_r-y_r)+\sum_{r\in\mathcal{R}_{\mathcal{U}}\cup\mathcal{R}^c_{\mathcal{U}}}\kappa_r(x,\varphi(x))\rho_c(y'_r-y_r).$$

By definition of $\widetilde{\mathcal{R}}_1$ and $\widetilde{\kappa}_{\widetilde{r}}(x)$ for $\widetilde{r}\in\widetilde{\mathcal{R}}_1$, the first summand of $f(x)$ and that of $\widetilde{g}(x)$ agree. Therefore, using the definition of $\widetilde{\mathcal{R}}_2$, all we need to prove is that

$$\sum_{\sigma\in\widetilde{\Delta}}q(x)\Pi(\sigma)\sum_{e\in\sigma}\rho_c(y'_{r(e)}-y_{r(e)})=\sum_{r\in\mathcal{R}_{\mathcal{U}}\cup\mathcal{R}^c_{\mathcal{U}}}\kappa_r(x,\varphi(x))\rho_c(y'_r-y_r).\tag{15}$$

Given a reaction $r\in\mathcal{R}_{\mathcal{U}}\cup\mathcal{R}^c_{\mathcal{U},y}$, the unique species in $\mathcal{U}$ that appears in the reactant is $U_i=s(e(r))$. Since the kinetics κ is $\mathcal{U}$-linear, we have

$$\kappa_r(x,\varphi(x))=\begin{cases}\varphi_i(x)v_r(x) & \text{if } r\in\mathcal{R}_{\mathcal{U}}\cup\mathcal{R}^c_{\mathcal{U},y}\\ v_r(x) & \text{if } r\in\mathcal{R}^c_{\mathcal{U},y'}.\end{cases}$$

For $r \in \mathcal{R}^c_{\mathcal{U},y'}$, we have $s(e(r)) = *$. Using the definition of $q(x)$ in (10) and that of $\varphi_i(x)$ in (11), we conclude that

$$\kappa_r(x, \varphi(x)) = q(x) \left(\sum_{\tau \in \Theta(s(e(r)))} \pi(\tau) \right) v_r(x),$$

for all $r \in \mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U}}$. Comparing this equality to (15), the statement follows if the following holds:

$$\sum_{r \in \mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U}}} \sum_{\tau \in \Theta(s(e(r)))} \pi(\tau) v_r(x) \rho_c(y'_r - y_r) = \tag{16}$$
$$\sum_{\sigma \in \widetilde{\Delta}} \sum_{e \in \sigma} \Pi(\sigma) \rho_c(y'_{r(e)} - y_{r(e)}).$$

We show the equality from right to left. Since either $\sum_{e \in \sigma} \rho_c(y'_{r(e)} - y_{r(e)}) = 0$ or $\Pi(\sigma) = 0$ for $\sigma \in \Delta \setminus \widetilde{\Delta}$, we obtain that

$$\begin{aligned} \sum_{\sigma \in \widetilde{\Delta}} \sum_{e \in \sigma} \Pi(\sigma)(y'_{r(e)} - y_{r(e)}) &= \sum_{\sigma \in \Delta} \sum_{e \in \sigma} \Pi(\sigma)(y'_{r(e)} - y_{r(e)}) \\ &= \sum_{\sigma \in \Delta} \sum_{e \in \sigma} \sum_{\gamma \in \Gamma(\sigma)} \pi(\gamma \setminus e) \pi(e)(y'_{r(e)} - y_{r(e)}) = (\star). \end{aligned}$$

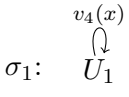
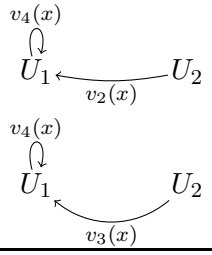
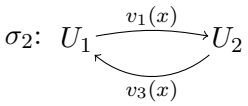
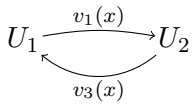
Each digraph $\gamma \setminus e$ in the sum is a spanning tree rooted at $s(e)$. There is a bijection between the set of triplets (σ, e, γ) such that $\sigma \in \Delta, e \in \sigma, \gamma \in \Gamma(\sigma)$ and the set of pairs (e, τ) such that e is an edge of $\mathcal{G}_{\mathcal{U}}$ and τ is a spanning tree rooted at $s(e)$. Using further that $\pi(e) = v_{r(e)}(x)$ and the correspondence (8), we obtain:

$$\begin{aligned} (\star) &= \sum_{e \in \mathcal{G}_{\mathcal{U}}} \sum_{\tau \in \Theta(s(e))} \pi(\tau) \pi(e)(y'_{r(e)} - y_{r(e)}) \\ &= \sum_{r \in \mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U}}} \sum_{\tau \in \Theta(s(e(r)))} \pi(\tau) v_r(x)(y'_r - y_r). \end{aligned}$$

This shows that (16) holds, which concludes the proof. □

Remark 25. We have defined a reaction network as a multigraph (Definition 3) to simplify the presentation, in particular in connection with Theorem 24, and hence also allowed multiple reactions between the same complexes. The reduced reaction network might, however, be further simplified by collapsing multiple reactions into one reaction without changing the ODE system, that is, using $\widehat{\mathcal{G}}$ as explained in Remark 8. The rate function of a new reaction is given by the sum of the rate functions being joined. See also Example 29.

Example 26. Let us consider Example 5 with $\mathcal{U}$ and $\mathcal{G}_{\mathcal{U}}$ given in Example 16. The set Δ has three elements, $\widetilde{\Delta}$ has two elements and Γ has four elements. The cycles in $\widetilde{\Delta}$ define two reactions in the reduced reaction network obtained by elimination of $\mathcal{U}$:

Cycle	$\Gamma(\sigma_i)$	**Reaction and Rate function**
σ_1: 		$S_2 \longrightarrow S_3$ $q(x)(v_2(x)+v_3(x))v_4(x)$
σ_2: 		$S_1 \longrightarrow S_2$ $q(x)v_1(x)v_3(x)$

Hence, the reduced reaction network is

$$S_1 \xrightarrow{q(x)v_1(x)v_3(x)} S_2 \xrightarrow{q(x)(v_2(x)+v_3(x))v_4(x)} S_3.$$

Procedure: We give a procedure to find the reactions in $\widetilde{\mathcal{R}}_2$ in Definition 23:

(i) Consider a cycle σ in $\mathcal{G}_{\mathcal{U}}$ with $\Gamma(\sigma) \neq \emptyset$ and such that

$$\sum_{e\in\sigma} \rho_c(y'_{r(e)} - y_{r(e)}) \neq 0.$$

(ii) List the reactions corresponding to the edges in the cycle.

(iii) Remove the eliminated species from the reactants and the products.

(iv) Sum the remaining species in the reactants and the products to obtain the reactant and the product of the new reaction.

(v) The rate function for the new reaction is $q(x)\Pi(\sigma)$ (as in (9), (10) and (14)).

To illustrate the method, we apply it to the cycle σ_2 in Example 26:

$$\begin{array}{rrcl} r_1\colon & S_1 + \cancel{U_1} & \longrightarrow & \cancel{U_2} \\ r_3\colon & \cancel{U_2} & \longrightarrow & \cancel{U_1} + S_2 \\ \hline & S_1 & \longrightarrow & S_2 \end{array}$$

Remark 27. Let $\sigma \in \Delta$. Each node U_i of σ is the source of exactly one edge and the target of exactly one edge of σ. Moreover, since $\mathcal{U}$ is a noninteracting set, each reaction involves at most one species in $\mathcal{U}$ in the reactant and one in the product. Therefore,

$$\sum_{e\in\sigma} \rho(y'_{r(e)} - y_{r(e)}) = 0.$$

In particular, each species that is a node of the cycle is removed exactly once as a reactant and once as a product in step (iii) of the procedure.

Remark 28. Let $r_{i_1}, \dots, r_{i_l} \in \mathcal{R}$ be reactions in a reaction network $(\mathcal{C}, \mathcal{R})$ that form a cycle. Then,

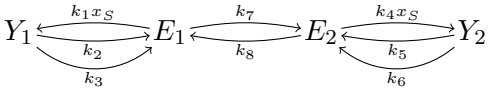
$$\sum_{j=1}^{l} \left(y'_{r_{i_j}} - y_{r_{i_j}} \right) = 0.$$

Therefore, a cycle in $\mathcal{G}_{\mathcal{U}}$ that corresponds to a cycle in $(\mathcal{C}, \mathcal{R})$ cannot define a reaction in $(\widetilde{\mathcal{C}}, \widetilde{\mathcal{R}})$, since it does not belong to $\widetilde{\Delta}$. If we nonetheless applied the procedure, the reaction would have the same reactant as product. In particular this is the case for any reversible reaction $\eta_1 \rightleftharpoons \eta_2$ with $\eta_1, \eta_2 \in \mathcal{C}$.

Example 29. Consider a phosphorylation mechanism for a substrate S such that phosphorylation may be catalyzed by two different forms of the same enzyme, E_1 and E_2, with mass-action kinetics:

$$S + E_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} Y_1 \xrightarrow{k_3} S_p + E_1 \qquad E_1 \underset{k_8}{\overset{k_7}{\rightleftharpoons}} E_2$$
$$S + E_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} Y_2 \xrightarrow{k_6} S_p + E_2$$

We consider the cut $\mathcal{U} = \{E_1, E_2, Y_1, Y_2\}$ and the reduced reaction network obtained by elimination of the cut. The corresponding multidigraph $\mathcal{G}_{\mathcal{U}}$ is



The set $\widetilde{\Delta}$ consists of two cycles

$$\sigma_1 : \; E_1 \underset{k_3}{\overset{k_1 x_S}{\rightleftharpoons}} Y_1 \qquad \sigma_2 : \; E_2 \underset{k_6}{\overset{k_4 x_S}{\rightleftharpoons}} Y_2$$

giving, by Theorem 24, the two reactions

$$S \xrightarrow{\widetilde{\kappa}_1(x)} S_p \qquad \text{with } \widetilde{\kappa}_1(x) = q(x) k_1 k_3 k_8 (k_5 + k_6) x_S$$
$$S \xrightarrow{\widetilde{\kappa}_2(x)} S_p \qquad \text{with } \widetilde{\kappa}_2(x) = q(x) k_4 k_6 k_7 (k_2 + k_3) x_S$$

and $q(x) = T_E \left((k_1 x_S + k_2 + k_3)(k_5 + k_6) k_8 + (k_2 + k_3)(k_4 x_S + k_5 + k_6) k_7 \right)^{-1}$.

The two cycles define the same reaction. Hence, by Remark 25, we might further simplify the network to

$$S \xrightarrow{\widetilde{\kappa}_1(x) + \widetilde{\kappa}_2(x)} S_p.$$

4.3 Elimination, reduction and connected components of $\mathcal{G}_{\mathcal{U}}$

In this section we consider the case where $\mathcal{G}_{\mathcal{U}}$ consists of several connected components.

Let $\mathcal{U} = \mathcal{U}_1 \sqcup \cdots \sqcup \mathcal{U}_l$ be the disjoint union of l sets of species such that each $\mathcal{G}_{\mathcal{U}_i}$ is a connected component of $\mathcal{G}_{\mathcal{U}}$. Assume that the kinetics is $\mathcal{U}$-linear and that assumption (ST) holds for $\mathcal{G}_{\mathcal{U}_i}$, $i = 1, \dots, l$.

Since there is at most one node $*$ in $\mathcal{G}_{\mathcal{U}}$, then there is at most one component that contains $*$. If $\mathcal{U}$ is not a cut, let $\mathcal{G}_{\mathcal{U}_l}$ be the component that contains $*$. If $\mathcal{U}$ is a cut, we assume for

convenience that $\mathcal{U}_l = \emptyset$. It follows that the sets $\mathcal{U}_1, \ldots, \mathcal{U}_{l-1}$ are cuts. Hence, given a non-empty stoichiometric compatibility class, the following equations hold:

$$\sum_{j \,|\, U_j \in \mathcal{U}_i} u_j = T_i, \quad T_i \in \mathbb{R}_{\geq 0} \qquad \text{for } i = 1, \ldots l-1. \tag{17}$$

The system of equations defined by $\dot{u}_1 = \cdots = \dot{u}_m = 0$ together with the equations in (17) is a linear system in $u_1, \ldots, u_m$. Since $\mathcal{U}$ is noninteracting and the kinetics is $\mathcal{U}$-linear, the system decomposes into l linear systems that can be solved separately:

$$\begin{cases} \dot{u}_j = 0 & \text{if } U_j \in \mathcal{U}_i \\ \sum\limits_{j \,|\, U_j \in \mathcal{U}_i} u_j = T_i & \text{if } i < l, \end{cases} \tag{18}$$

for $i = 1, \ldots, l$. For each i, this system is precisely the system of equations solved for the elimination of the species in $\mathcal{U}_i$. Theorem 17 holds in each connected component $\mathcal{G}_{\mathcal{U}_i}$ and hence elimination of $\mathcal{U}$ can be done by eliminating each subset $\mathcal{U}_i$ separately. In particular, each component defines a function $q_i(x)$, as in (9) or (10). We let $\widetilde{\Omega} \subset \Omega$ be the domain where the functions $q_i(x)$, for $i = 1, \ldots, l$, are all defined.

For $i = 1, \ldots, l$, the elimination procedure defines functions

$$\varphi^i(x) = (\varphi^i_1(x), \ldots, \varphi^i_{m_i}(x)),$$

where m_i is the number of species in $\mathcal{U}_i$, such that the equations in (18) hold if and only if

$$u_j = \varphi^i_{j_i}(x) \quad \text{if } U_j \text{ is the } j_i\text{th species in } \mathcal{U}_i.$$

By possibly reordering the set of species $\mathcal{U}$, the function

$$\varphi(x) = (\varphi^1(x), \ldots, \varphi^l(x))$$

fulfils that

$$u_j = \varphi_j(x)$$

if and only if the system $\dot{u}_1 = \cdots = \dot{u}_m = 0$ together with the equations in (17) holds.

Let Γ_i be the set of sub-multidigraphs defined in (13) for the graph $\mathcal{G}_{\mathcal{U}_i}$. Then, the set Γ for $\mathcal{G}_{\mathcal{U}}$ is defined as

$$\Gamma = \Gamma_1 \cup \cdots \cup \Gamma_l.$$

For σ a cycle in $\mathcal{G}_{\mathcal{U}_i}$, the set $\Gamma(\sigma)$ is computed over the elements in Γ_i, that is, $\Gamma(\sigma) \subseteq \Gamma_i$. Hence, the sets of cycles Δ and $\widetilde{\Delta}$ of $\mathcal{G}_{\mathcal{U}}$ decompose also in l subsets according to the connected component that contains each cycle.

The reduced reaction network $(\widetilde{\mathcal{C}}, \widetilde{\mathcal{R}})$ obtained by elimination of $\mathcal{U}$ is defined as in Definition 23 using the function φ and the domain $\widetilde{\Omega}$. The rate function for a reaction $\widetilde{r}_\sigma$ for a cycle σ in $\mathcal{G}_{\mathcal{U}_i}$ is $\widetilde{\kappa}_{\widetilde{r}_\sigma} = q_i(x)\Pi(\sigma)$. Since $\Gamma(\sigma)$ is computed over the elements in Γ_i, so is $\Pi(\sigma)$. The set of reactions $\widetilde{\mathcal{R}}_2$ can be decomposed as well into l subsets according to the decomposition of $\widetilde{\Delta}$.

We might further simplify the elimination of the species in $\mathcal{U}_l$, and hence the computation of the reactions in $\widetilde{\mathcal{R}}_2$ defined by cycles in $\mathcal{G}_{\mathcal{U}_l}$. We let $\mathcal{G}^*_{\mathcal{U}_l}$ be the multidigraph obtained by removal of the node $*$ and all the edges that have $*$ as source or target node.

Proposition 30. *Let $\mathcal{U}$ be a noninteracting set that is not a cut and assume that $\mathcal{G}_{\mathcal{U}}$ is connected. Let $\mathcal{H} \subseteq \mathcal{U}$ be the set of nodes of a connected component of $\mathcal{G}_{\mathcal{U}}^*$ and let $\Theta_{\mathcal{H}}(N)$ be the set of spanning trees rooted at $N \in \mathcal{H} \cup \{*\}$ in the sub-multidigraph $\mathcal{G}_{\mathcal{H}}$ of $\mathcal{G}_{\mathcal{U}}$. Then,*

$$q(x) \sum_{\tau \in \Theta(N)} \pi(\tau) = \frac{\sum\limits_{\tau \in \Theta(N)} \pi(\tau)}{\sum\limits_{\tau \in \Theta(*)} \pi(\tau)} = \frac{\sum\limits_{\tau \in \Theta_{\mathcal{H}}(N)} \pi(\tau)}{\sum\limits_{\tau \in \Theta_{\mathcal{H}}(*)} \pi(\tau)} = q_{\mathcal{H}}(x) \sum_{\tau \in \Theta_{\mathcal{H}}(N)} \pi(\tau)$$

with $q_{\mathcal{H}}(x)$ as in (10) *for the sub-multidigraph $\mathcal{G}_{\mathcal{H}}$ of $\mathcal{G}_{\mathcal{U}}$.*

Proof. Note that $\mathcal{H}$ and $\mathcal{U} \setminus \mathcal{H}$ are noninteracting sets, hence $\mathcal{G}_{\mathcal{H}}$ and $\mathcal{G}_{\mathcal{U}\setminus\mathcal{H}}$ are defined. Let $\Theta_{\mathcal{G}_{\mathcal{U}\setminus\mathcal{H}}}(*)$ be the set of spanning trees of $\mathcal{G}_{\mathcal{U}\setminus\mathcal{H}}$ rooted at $*$. Let $\tau \in \Theta(N)$ be a spanning tree of $\mathcal{G}_{\mathcal{U}}$ rooted at N. The path from any node $N' \notin \mathcal{H}$ to N contains $*$ by hypothesis. Therefore, τ is the union of a spanning tree τ_1 of $\mathcal{G}_{\mathcal{H}}$ rooted at N and a spanning tree τ_2 of $\mathcal{G}_{\mathcal{U}\setminus\mathcal{H}}$ rooted at $*$. Reciprocally, the union of any pair of spanning trees $\tau_1 \in \Theta_{\mathcal{H}}(N)$, $\tau_2 \in \Theta_{\mathcal{G}_{\mathcal{U}}\setminus\mathcal{H}}(*)$ is a spanning tree of $\mathcal{G}_{\mathcal{U}}$ rooted at N. As subgraphs of $\mathcal{G}_{\mathcal{U}}$, τ_1 and τ_2 intersect at $*$. Hence $\pi(\tau) = \pi(\tau_1)\pi(\tau_2)$ and we obtain

$$\sum_{\tau \in \Theta(N)} \pi(\tau) = \left(\sum_{\tau_1 \in \Theta_{\mathcal{H}}(N)} \pi(\tau_1) \right) \left(\sum_{\tau_2 \in \Theta_{\mathcal{G}_{\mathcal{U}}\setminus\mathcal{H}}(*)} \pi(\tau_2) \right),$$

from where it follows that:

$$\frac{\sum\limits_{\tau \in \Theta(N)} \pi(\tau)}{\sum\limits_{\tau \in \Theta(*)} \pi(\tau)} = \frac{\left(\sum\limits_{\tau_1 \in \Theta_{\mathcal{H}}(N)} \pi(\tau_1) \right) \left(\sum\limits_{\tau_2 \in \Theta_{\mathcal{G}_{\mathcal{U}}\setminus\mathcal{H}}(*)} \pi(\tau_2) \right)}{\left(\sum\limits_{\tau_1 \in \Theta_{\mathcal{H}}(*)} \pi(\tau_1) \right) \left(\sum\limits_{\tau_2 \in \Theta_{\mathcal{G}_{\mathcal{U}}\setminus\mathcal{H}}(*)} \pi(\tau_2) \right)} = \frac{\sum\limits_{\tau \in \Theta_{\mathcal{H}}(N)} \pi(\tau)}{\sum\limits_{\tau \in \Theta_{\mathcal{H}}(*)} \pi(\tau)}.$$

□

Proposition 30 guarantees that the computation of $\varphi^l(x)$ and the rate functions $q_l(x)\Pi(\sigma)$ of the reduced reaction network in Definition 23 is independent of whether it is performed considering $\mathcal{G}_{\mathcal{U}_l}$ or $\mathcal{G}_{\mathcal{H}_i}$, with $\mathcal{H}_i \subset \mathcal{U}_l$ being the set of nodes in a connected component of $\mathcal{G}_{\mathcal{U}_l}^*$.

As an illustration of this procedure, see Example 39 in Subsection 6.2.

4.4 Rate functions and kinetics

In this section we study the kinetics of the reduced reaction network in relation to the kinetics of the original reaction network. As noted in Remark 9, a biochemically meaningful rate function fulfils $\kappa_r(x) = 0$ whenever the concentration of one of the reactant species is zero, that is, $\operatorname{supp}(y_r) \not\subseteq \operatorname{supp}(x)$ implies $\kappa_r(x) = 0$.

Based on this, we say that a function f defined on a subset $V \subseteq \mathbb{R}^n$ is **standard** for $r \in \mathcal{R}$ if

$$\text{for all} \quad x \in V: \quad \operatorname{supp}(y_r) \not\subseteq \operatorname{supp}(x) \quad \Rightarrow \quad f(x) = 0. \tag{19}$$

We say that the function is **fully standard** if the reverse implication in (19) holds as well.

The rate function κ_r is (fully) standard if it is a (fully) standard function for r. Similarly, the kinetics κ is (fully) standard if the rate function of all $r \in \mathcal{R}$ is (fully) standard.

We show in Corollary 1 below that if $\kappa(x,u)$, defined on $\Omega\times\mathbb{R}^m_{\geq 0}$, is standard, then so is $\widetilde{\kappa}(x)$, defined on $\widetilde{\Omega}$. Recall that $\widetilde{\Omega}$ is the subset of Ω for which the denominator of $q(x)$ does not vanish. Therefore, there are no points in $\widetilde{\Omega}$ where both the numerator and the denominator of $\widetilde{\kappa}_{\widetilde{r}}$ vanish for some $\widetilde{r}\in\widetilde{\mathcal{R}}$.

Proposition 31. *Let $\sigma\in\widetilde{\Delta}$ and consider the reaction $\widetilde{r}_\sigma\in\widetilde{\mathcal{R}}_2$. If the kinetics $\kappa(x,u)$ is standard (resp. fully standard) on $\Omega\times\mathbb{R}^m_{\geq 0}$, then $\pi(\sigma)$ is standard (resp. fully standard) on $\widetilde{\Omega}$ for $\widetilde{r}_\sigma$.*

Proof. By definition

$$\pi(\sigma)=\prod_{e\in\sigma}v_{r(e)}(x),\qquad y_{\widetilde{r}_\sigma}=\sum_{e\in\sigma}\rho_c(y_{r(e)}).$$

Hence, $\pi(\sigma)=0$ if and only if $v_{r(e)}(x)=0$ for some $e\in\sigma$. Further, we have that $\operatorname{supp}(y_{\widetilde{r}_\sigma})=\bigcup_{e\in\sigma}\operatorname{supp}(y_{r(e)})$.

Assume that the kinetics κ is standard on $\Omega\times\mathbb{R}^m_{\geq 0}$. For $x\in\widetilde{\Omega}$, we have

$$\begin{aligned}\operatorname{supp}(y_{\widetilde{r}_\sigma})\not\subseteq\operatorname{supp}(x)&\Leftrightarrow\exists\, e\in\sigma\colon\ \operatorname{supp}(\rho_c(y_{r(e)}))\not\subseteq\operatorname{supp}(x)\\&\Rightarrow\exists\, e\in\sigma\colon v_{r(e)}(x)=0\\&\Leftrightarrow\pi(\sigma)=0.\end{aligned}$$

This shows that $\pi(\sigma)$ is standard on $\widetilde{\Omega}$ for $\widetilde{r}_\sigma$. If the kinetics is fully standard, then the reverse of the second implication holds, showing that $\pi(\sigma)$ is also fully standard for $\widetilde{r}_\sigma$. □

Corollary 1. *If the kinetics $\kappa(x,u)$ is standard on $\Omega\times\mathbb{R}^m_{\geq 0}$, then so is $\widetilde{\kappa}(x)$ on $\widetilde{\Omega}$.*

Proof. For $\widetilde{r}_\sigma\in\widetilde{\mathcal{R}}_2$ defined by $\sigma\in\widetilde{\Delta}$, we have $\widetilde{\kappa}_{\widetilde{r}_\sigma}(x)=q(x)\Pi(\sigma)$. By equation (14), $\pi(\sigma)$ is a factor of $\Pi(\sigma)$ and hence $\widetilde{\kappa}_{\widetilde{r}_\sigma}$ is standard on $\widetilde{\Omega}$ by Proposition 31.

For $\widetilde{r}=\rho_c(y_r)\to\rho_c(y_r')\in\widetilde{\mathcal{R}}_1$ with $r\in\mathcal{R}_{\mathcal{U}}^{\perp}$, we have $\widetilde{\kappa}_{\widetilde{r}}(x)=\kappa_r(x,\varphi(x))$ by Definition 23. Since $r\in\mathcal{R}_{\mathcal{U}}^{\perp}$, then $\operatorname{supp}(y_r)=\operatorname{supp}(\rho_c(y_r))=\operatorname{supp}(y_{\widetilde{r}})$. Since κ_r is standard on $\Omega\times\mathbb{R}^m_{\geq 0}$, then $\kappa_r(x,u)$ vanishes if $x\in\Omega$ fulfils $\operatorname{supp}(y_{\widetilde{r}})\not\subseteq\operatorname{supp}(x)$. Since the denominator of $q(x)$ does not vanish for $x\in\widetilde{\Omega}$, we have $\widetilde{\kappa}_{\widetilde{r}}(x)=0$ whenever $\operatorname{supp}(y_{\widetilde{r}})\not\subseteq\operatorname{supp}(x)$ and $x\in\widetilde{\Omega}$, showing that $\widetilde{\kappa}_{\widetilde{r}}(x)$ is standard on $\widetilde{\Omega}$. □

If the original reaction network is endowed with mass-action kinetics we have a sharper result than Corollary 1, as mass-action kinetics are fully standard:

Proposition 32. *Assume that the reaction network $(\mathcal{C},\mathcal{R})$ is endowed with mass-action kinetics. For $\sigma\in\widetilde{\Delta}$, there exists a constant $k_{\widetilde{r}_\sigma}>0$ such that*

$$\pi(\sigma)=k_{\widetilde{r}_\sigma}x^{y_{\widetilde{r}_\sigma}},\quad x\in\widetilde{\Omega}.$$

Proof. The label of an edge of σ is $k_{r(e)}x^{\rho_c(y_{r(e)})}$, where $k_{r(e)}>0$ is the reaction rate constant of reaction $r(e)$. Using the definition of $y_{\widetilde{r}_\sigma}$ in Definition 23, we obtain

$$\pi(\sigma)=\prod_{e\in\sigma}k_{r(e)}x^{\rho_c(y_{r(e)})}=\left(\prod_{e\in\sigma}k_{r(e)}\right)x^{\sum_{e\in\sigma}\rho_c(y_{r(e)})}=\left(\prod_{e\in\sigma}k_{r(e)}\right)x^{y_{\widetilde{r}_\sigma}}.$$

Hence $\pi(\sigma)$ has the claimed form with $k_{\widetilde{r}_\sigma}=\prod_{e\in\sigma}k_{r(e)}$. □

The following corollary follows from the fact that $\pi(\sigma)$ is a factor of $\Pi(\sigma)$, see equation (14).

Corollary 2. *Assume that the reaction network $(\mathcal{C}, \mathcal{R})$ is endowed with mass-action kinetics. If a species S_i is involved in the reactant of $\widetilde{r}_\sigma$, then its concentration x_i is a factor of $\Pi(\sigma)$ with exponent at least $(y_{\widetilde{r}_\sigma})_i$.*

The kinetics of the reduced reaction network might not be fully standard, even if the kinetics of the original network is fully standard. For a reaction $\widetilde{r} \in \widetilde{\mathcal{R}}_1$ defined by $r \in \mathcal{R}_{\mathcal{U}}^{\perp}$, the rate function $\widetilde{\kappa}_{\widetilde{r}}(x)$ is fully standard on $\widetilde{\Omega}$ provided $\kappa_r(x,u)$ is on $\Omega \times \mathbb{R}^m_{\geq 0}$. However, different situations might occur for reactions $\widetilde{r}_\sigma \in \widetilde{\mathcal{R}}_2$, where the rate function might vanish if

$$\sum_{\gamma \in \Gamma(\sigma)} \pi(\gamma \setminus \sigma) = 0 \quad \text{for some} \quad x \in \widetilde{\Omega}$$

(cf. equation (14)). We characterise when this happens in Lemma 3 below.

Under mass-action kinetics, it might also be the case that the exponents of the concentrations of the species involved in the reactant of $\widetilde{r}_\sigma$ are larger than their stoichiometric coefficients in the reaction $\widetilde{r}_\sigma$. Also, there may be factors that depend on the concentration of species not involved in the reactant. See Example 33.

When applying Remark 25, it may happen that the rate of the reaction becomes fully standard after joining two reactions corresponding to different cycles. See Example 33.

Lemma 3. *Assume that $\kappa(x,u)$ is fully standard on $\Omega \times \mathbb{R}^m_{\geq 0}$. Let $\sigma \in \widetilde{\Delta}$ and $x \in \widetilde{\Omega}$ be such that $\operatorname{supp}(y_{\widetilde{r}_\sigma}) \subseteq \operatorname{supp}(x)$. Then the following two statements are equivalent:*

(i) $\Pi(\sigma)$ vanishes at x.

(ii) $\varphi_j(x) = 0$, for all $j \in \{1, \dots, m\}$ such that $U_j \in \mathcal{U}$ is a node in σ.

Proof. Throughout the proof all functions are assumed to be evaluated at x.

Assume that statement (ii) holds. Let $e \in \sigma$ be an edge and let $U_j = s(e)$. Then

$$\Pi(\sigma) = \pi(e) \sum_{\gamma \in \Gamma(\sigma)} \pi(\gamma \setminus e) \leq \pi(e) \sum_{\tau \in \Theta(U_j)} \pi(\tau) = \pi(e) \frac{\varphi_j(x)}{q(x)}.$$

Since $\varphi_j(x) = 0$, $\Pi(\sigma) = 0$ and statement (i) holds.

Assume now that statement (i) holds. Let $U_j \in \mathcal{U}$ be a node in σ, τ a spanning tree rooted at U_j and $e \in \sigma$ the edge with source U_j. We construct a new tree $\widehat{\tau}$ as follows: for every $U_l \neq U_j$ in σ, replace the only edge with source U_l in the tree τ by the edge in σ with source node U_l. The obtained subgraph $\widehat{\tau}$ is also a spanning tree rooted at U_j that satisfies $\widehat{\tau} \cup e \in \Gamma(\sigma)$.

By assumption (i) and the definition of $\Pi(\sigma)$, $\pi(\widehat{\tau} \cup e) = 0$. Since $\kappa(x,u)$ is fully standard, $\pi(e') \neq 0$ for all $e' \in \sigma$ by Proposition 31. Therefore there must be an edge $\widehat{e}$ of $\widehat{\tau}$ that does not belong to σ and such that $\pi(\widehat{e}) = 0$. By construction, this edge is also an edge of τ and hence $\pi(\tau) = 0$.

This proves that $\pi(\tau) = 0$ for all $\tau \in \Theta(U_j)$, which implies that $\varphi_j(x) = 0$ (cf. (11)). Hence statement (ii) holds. ∎

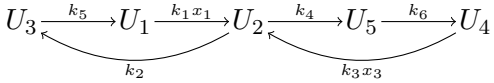
The lemma states that the rate function for $\widetilde{r}_\sigma$ vanishes if $\kappa_{r(e)}(x, \varphi(x)) = 0$ for all $e \in \sigma$. This holds because the kinetics is $\mathcal{U}$-linear and hence u_j is a factor of the rate function of the reaction $r = r(e)$ with $s(e) = U_j$.

In some cases, the points $x \in \widetilde{\Omega}$ for which Lemma 3 applies are precisely those that satisfy $x_i = 0$ for a certain $i \in \{1, \ldots, p\}$. If that is so then the reaction can be modified by adding S_i to the reactant and the product of $\widetilde{r}_\sigma$, so that the rate function $\widetilde{\kappa}_{\widetilde{r}_\sigma}$ is fully standard for the modified reaction. The ODE system associated with this reduced reaction network is the same. We have chosen to omit modifying reactions in this way.

Example 33. Consider the following reaction network with mass-action kinetics:

$$\begin{array}{lll} S_1 + U_1 \xrightarrow{k_1} U_2 \xrightarrow{k_2} S_2 + U_3 & U_3 \xrightarrow{k_5} U_1 & S_2 \xrightarrow{k_7} S_1 \\ S_3 + U_4 \xrightarrow{k_3} U_2 \xrightarrow{k_4} S_4 + U_5 & U_5 \xrightarrow{k_6} U_4 & S_4 \xrightarrow{k_8} S_3 \end{array} \tag{20}$$

and the cut $\mathcal{U} = \{U_1, U_2, U_3, U_4, U_5\}$. The multidigraph $\mathcal{G}_\mathcal{U}$ is



The reduced reaction network obtained by elimination of $\mathcal{U}$ is

$$S_1 \underset{\widetilde{\kappa}_2(x)}{\overset{\widetilde{\kappa}_1(x)}{\rightleftharpoons}} S_2 \qquad S_3 \underset{\widetilde{\kappa}_4(x)}{\overset{\widetilde{\kappa}_3(x)}{\rightleftharpoons}} S_4 \tag{21}$$

where

$$\begin{array}{ll} \widetilde{\kappa}_1(x) = q(x)k_1k_2k_3k_5k_6x_1\boldsymbol{x_3}, & \widetilde{\kappa}_2(x) = k_7x_2 \\ \widetilde{\kappa}_3(x) = q(x)k_1k_3k_4k_5k_6\boldsymbol{x_1}x_3, & \widetilde{\kappa}_4(x) = k_8x_4 \end{array}$$

and

$$q(x) = T\left((k_2k_6 + k_4k_5 + k_5k_6)k_1k_3x_1x_3 + k_1k_4k_5k_6x_1 + k_2k_3k_5k_6x_3\right)^{-1}.$$

Concentrations in bold correspond to species that are not in the reactant of the corresponding reaction.

We have $\widetilde{\Omega} = \mathbb{R}^4_{\geq 0} \backslash \{x : x_1 = x_3 = 0\}$. The species S_3 does not appear in the reactant of the first reaction in the reduced reaction network, but its rate function vanishes for $x_3 = 0$. This reaction corresponds to the left-cycle of the multidigraph $\mathcal{G}_\mathcal{U}$, which involves nodes U_1, U_2, U_3. By Lemma 3, the functions $\varphi_1(x), \varphi_2(x), \varphi_3(x)$ also vanish when $x_3 = 0$. Indeed:

$$\begin{array}{ll} \varphi_1(x) = q(x)k_2k_3k_5k_6\boldsymbol{x_3}, & \varphi_2(x) = q(x)k_1k_3k_5k_6x_1\boldsymbol{x_3}, \\ \varphi_3(x) = q(x)k_1k_2k_3k_6x_1\boldsymbol{x_3}. & \end{array}$$

Consider the reaction network (20) with an additional reaction:

$$S_1 + U_4 \xrightarrow{k_9} S_2 + U_4.$$

The set of species $\mathcal{U} = \{U_1, U_2, U_3, U_4, U_5\}$ is again a cut for the extended reaction network. The reaction r_9 defines a self-edge for U_4 in $\mathcal{G}_\mathcal{U}$, that defines an element of $\widetilde{\Delta}$. Since $e(r_9)$ is a self-edge, neither the set of spanning trees, nor the reactions defined by the other elements in $\widetilde{\Delta}$, nor their rate functions are modified with respect to the above computations. The cycle given by $e(r_9)$ defines a reaction $S_1 \longrightarrow S_2$ in $\widetilde{\mathcal{R}}_2$ with rate function $q(x)k_1k_4k_5k_6k_9\boldsymbol{x_1^2}$.

Since there are now two reactions $S_1 \longrightarrow S_2$ in $\widetilde{\mathcal{R}}_2$, they can be joined and thus the reduced reaction network is also (21) but now with rate functions:

$$\begin{array}{l} \widetilde{\kappa}_1(x) = q(x)k_1k_5k_6(k_2k_3x_1\boldsymbol{x_3} + k_4k_9\boldsymbol{x_1^2}) = q(x)k_1k_5k_6(\boldsymbol{k_2x_3 + k_4x_1})x_1, \\ \widetilde{\kappa}_2(x) = k_7x_2 \quad \widetilde{\kappa}_3(x) = q(x)k_1k_3k_4k_5k_6\boldsymbol{x_1}x_3, \quad \widetilde{\kappa}_4(x) = k_8x_4 \end{array} \tag{22}$$

and $q(x)$ as above. The kinetics (22) is fully standard for the reaction network (21). Factors in bold correspond to species that are not in the reactant of the corresponding reaction or have an exponent larger than the corresponding stoichiometric coefficient.

5 Conservation laws of the two networks

Consider a reaction network $(\mathcal{C}, \mathcal{R})$ on $\mathcal{S}$, a set of noninteracting species $\mathcal{U} \subseteq \mathcal{S}$ and a $\mathcal{U}$-linear kinetics κ. Further assume that $\mathcal{G}_{\mathcal{U}}$ is connected and that assumption (ST) holds. Let $(\widetilde{\mathcal{C}}, \widetilde{\mathcal{R}})$ be the reduced reaction network on $\mathcal{U}^c$ obtained by elimination of $\mathcal{U}$ and let $\widetilde{S} \subseteq \mathbb{R}^p$ be its stoichiometric subspace.

In this section we compare the spaces $S^\perp$ and $\widetilde{S}^\perp$, which define the conservation laws of each reaction network. Let $\overline{\ell} = |\mathcal{R}_{\mathcal{U}} \cup \mathcal{R}^c_{\mathcal{U}}|$.

For this, we need some additional graphical constructions. Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be a multi-digraph and assume that the sets $\mathcal{N} = \{N_1, \ldots, N_n\}$ and $\mathcal{E} = \{e_1, \ldots, e_\ell\}$ are ordered. We define the following objects:

1. The **incidence matrix** $C_{\mathcal{G}}$ of $\mathcal{G}$ is the $n \times \ell$ real matrix such that

$$(C_{\mathcal{G}})_{ij} = \begin{cases} 1 & \text{if} \quad N_i = t(e_j) \neq s(e_j) \\ -1 & \text{if} \quad N_i = s(e_j) \neq t(e_j) \\ 0 & \text{otherwise.} \end{cases}$$

2. The **cycle space** of $\mathcal{G}$ is the kernel of the incidence matrix.

If $\sigma \in \Delta$, that is, if σ is a cycle of $\mathcal{G}$, then the vector ν_σ with $(\nu_\sigma)_i = 1$ if e_i is an edge of the cycle and $(\nu_\sigma)_i = 0$ otherwise, belongs to the cycle space of $\mathcal{G}$. Moreover, the elements ν_σ correspond to the irreducible elements in the cycle space with all non-zero components equal to one. That is, ν_σ cannot be expressed as the positive sum of two vectors in the cycle space with non-negative integer coordinates. The elements in Δ are also called elementary cycles in the literature ([19]).

Now let $C_{\mathcal{G}}$ be the incidence matrix of $\mathcal{G}_{\mathcal{U}}$ (assuming nodes and edges are ordered). Note that $C_{\mathcal{G}} \in \mathbb{R}^{\overline{m} \times \overline{\ell}}$ with $\overline{m} = m, m+1$ depending on whether $\mathcal{U}$ is a cut or not.

Assume that the set of cycles $\Delta = \{\sigma_1, \ldots, \sigma_{|\Delta|}\}$ is ordered, and let H be the $|\Delta| \times \overline{\ell}$ matrix whose ith row is ν_{σ_i}.

Lemma 4. *If $\mathcal{G}_{\mathcal{U}}$ is strongly connected, then* $\ker H = \operatorname{im} C^t_{\mathcal{G}}$.

Proof. By duality it is enough to show that $\operatorname{im} H^t = \ker C_{\mathcal{G}}$. Since for all $\sigma \in \Delta$, ν_σ belongs to the cycle space, we have $C_{\mathcal{G}} H^t = 0$, that is, $\operatorname{im} H^t \subseteq \ker C_{\mathcal{G}}$. Since $\mathcal{G}_{\mathcal{U}}$ is strongly connected, then there exists a basis of $\ker C_{\mathcal{G}}$ whose elements are of the form ν_σ for $\sigma \in \Delta$ (see [1]). Hence $\operatorname{im} H^t \supseteq \ker C_{\mathcal{G}}$ as desired. □

Observe that the vector $\omega^0 = (0, \overset{p}{\ldots}, 0, 1, \overset{m}{\ldots}, 1)$ does not belong to $S^\perp$ unless $\mathcal{U}$ is a cut. With this observation we obtain the following theorem.

Theorem 34. *Let* $\omega^0 = (0, \overset{p}{\ldots}, 0, 1, \overset{m}{\ldots}, 1)$.

(i) If $\omega \in S^\perp$, *then* $\rho_c(\omega) \in \widetilde{S}^\perp$. *Moreover, if* $\omega^1, \ldots, \omega^h$ *are independent vectors in* $S^\perp$, *such that* $\omega^0 \notin \langle \omega^1, \ldots, \omega^h \rangle$, *then* $\rho_c(\omega^1), \ldots, \rho_c(\omega^h)$ *are independent vectors in* $\widetilde{S}^\perp$.

(ii) Assume that $\mathcal{G}_\mathcal{U}$ *is strongly connected and let* $\widetilde{\omega} \in \widetilde{S}^\perp$. *Then, there exists* $\omega \in S^\perp$ *such that* $\widetilde{\omega} = \rho_c(\omega)$.

Proof. (i) Let $\omega \in S^\perp$ and $\widetilde{r} \in \widetilde{\mathcal{R}}_1$ be a reaction corresponding to $r \in \mathcal{R}_\mathcal{U}^\perp$. Since $\omega \cdot (y'_r - y_r) = 0$ and $\rho(y'_r - y_r) = 0$, we deduce that $\rho_c(\omega) \cdot (y'_{\widetilde{r}} - y_{\widetilde{r}}) = 0$.

Let $\widetilde{r}_\sigma \in \widetilde{\mathcal{R}}_2$ for $\sigma \in \widetilde{\Delta}$ and $\eta = \sum_{e \in \sigma} (y'_{r(e)} - y_{r(e)})$. Then, by definition,

$$y'_{\widetilde{r}_\sigma} - y_{\widetilde{r}_\sigma} = \rho_c(\eta).$$

Using that $\rho(\eta) = 0$ (Remark 27), we have

$$\rho_c(\omega) \cdot (y'_{\widetilde{r}_\sigma} - y_{\widetilde{r}_\sigma}) = \rho_c(\omega) \cdot \rho_c(\eta) + \rho(\omega) \cdot \rho(\eta) = \omega \cdot \eta = 0.$$

This proves the first part of (i).

Let $\omega^1, \ldots, \omega^h \in S^\perp$ be as in the statement. Assume $\sum_{i=1}^h \lambda_i \rho_c(\omega^i) = 0$ for some $\lambda_1, \ldots, \lambda_h \in \mathbb{R}$ and let $\beta = \sum_{i=1}^h \lambda_i \omega^i \in S^\perp$. Since $\rho_c(\beta) = 0$, it follows that $\mathrm{supp}(\beta) \subseteq \mathcal{U}$. By Corollary 2, $\beta = 0$ if $\mathcal{U}$ is not a cut and $\beta = a\omega^0$ for some $a \in \mathbb{R}$ if $\mathcal{U}$ is a cut. Since by hypothesis $\omega^0 \notin \langle \omega^1, \ldots, \omega^h \rangle$, we have $a = 0$ and hence $\beta = 0$. It follows from the linear independence of $\omega^1, \ldots, \omega^h$, that $\lambda_1 = \cdots = \lambda_h = 0$. We have proven (i).

(ii) Assume that $\mathcal{G}_\mathcal{U}$ is strongly connected. Given $\widetilde{\omega} \in \widetilde{S}^\perp$, we want to prove that there exist $\alpha_1, \ldots, \alpha_m$ such that

$$\omega = (\widetilde{\omega}_1, \ldots, \widetilde{\omega}_p, \alpha_1, \ldots, \alpha_m) \in S^\perp.$$

That is, such that $\omega A = 0$, where A is the stoichiometric matrix of $(\mathcal{C}, \mathcal{R})$.

We order the set $\mathcal{R}$ in such a way that $\{r_{\overline{\ell}+1}, \ldots, r_\ell\} = \mathcal{R}_\mathcal{U}^\perp$. Then, A can be written in block form as

$$\begin{pmatrix} A_1 & A_3 \\ A_2 & 0_{m \times (\ell - \overline{\ell})} \end{pmatrix} \text{ with } A_1 \in \mathbb{R}^{p \times \overline{\ell}},\ A_2 \in \mathbb{R}^{m \times \overline{\ell}} \text{ and } A_3 \in \mathbb{R}^{p \times (\ell - \overline{\ell})}.$$

The columns of A_3 correspond to the reactions in $\widetilde{\mathcal{R}}_1$ and are thus vectors of $\widetilde{S}$ by Definition 23. Hence, $\omega \begin{pmatrix} A_3 \\ 0_{m \times (\ell - \overline{\ell})} \end{pmatrix} = 0$ for any choice of $\alpha_1, \ldots, \alpha_m$. Therefore,

$$\begin{aligned} \omega A = 0 &\Leftrightarrow (\widetilde{\omega}_1, \ldots, \widetilde{\omega}_p) A_1 + (\alpha_1, \ldots, \alpha_m) A_2 = 0 \\ &\Leftrightarrow A_2^t \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = - \begin{pmatrix} \widetilde{\omega} \cdot \rho_c(y'_{r_1} - y_{r_1}) \\ \vdots \\ \widetilde{\omega} \cdot \rho_c(y'_{r_{\overline{\ell}}} - y_{r_{\overline{\ell}}}) \end{pmatrix} =: v. \end{aligned} \tag{23}$$

Let H be as defined above Lemma 4. We argue next that $Hv = 0$, that is, $v \in \ker H$. Hence, by Lemma 4, $v \in \mathrm{im}\, C_\mathcal{G}^t$.

Indeed, if $\sigma \in \widetilde{\Delta}$, then $\widetilde{r}_\sigma \in \widetilde{\mathcal{R}}$ and hence

$$0 = \widetilde{\omega} \cdot (y'_{\widetilde{r}_\sigma} - y_{\widetilde{r}_\sigma}) = \sum_{e \in \sigma} \widetilde{\omega} \cdot \rho_c(y'_{r(e)} - y_{r(e)}).$$

Let $\sigma \in \Delta \setminus \widetilde{\Delta}$. Since $\mathcal{G}_\mathcal{U}$ is strongly connected, by Proposition 20 we deduce that $\Gamma(\sigma) \neq \emptyset$. Thus, by definition of $\widetilde{\Delta}$,

$$0 = \sum_{e \in \sigma} \rho_c(y'_{r(e)} - y_{r(e)}) \quad \text{and hence} \quad 0 = \sum_{e \in \sigma} \widetilde{\omega} \cdot \rho_c(y'_{r(e)} - y_{r(e)}).$$

If $\mathcal{U}$ is a cut, then $A_2 = C_\mathcal{G}$. Since $v \in \operatorname{im} C_\mathcal{G}^t$ we deduce that the system (23) has a solution $(\alpha_1, \ldots, \alpha_m)$.

If $\mathcal{U}$ is not a cut, then $A_2 = \widetilde{C}_\mathcal{G}$, with $\widetilde{C}_\mathcal{G}$ being the matrix obtained from $C_\mathcal{G}$ by removing the last row, corresponding to the node $*$. We can then rewrite our system of equations (23) as

$$v = A_2^t \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = C_\mathcal{G}^t \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \\ 0 \end{pmatrix}. \tag{24}$$

Since $v \in \operatorname{im} C_\mathcal{G}^t$, there exists a vector $b = (b_1, \ldots, b_{m+1})^t$ in $\mathbb{R}^{m+1}$ such that $C_\mathcal{G}^t b = v$. Since the column sums of $C_\mathcal{G}$ are all zero by definition, $(1, \ldots, 1)^t \in \ker C_\mathcal{G}^t$. Thus

$$v = C_\mathcal{G}^t(b - b_{m+1}(1, \ldots, 1)^t).$$

Therefore, (24) has a solution of the claimed form by setting $\alpha_i = b_i - b_{m+1}$, for $i = 1 \ldots, m$. □

Example 35. A minimal set of conservation laws for Example 5 is given in (7). A minimal set of conservation laws of the reduced reaction network in Example 26 contains the single element

$$x_1 + x_2 + x_3 = \widetilde{T},$$

which is the element obtained by deleting $u_1 = x_4$ and $u_2 = x_5$ from the first equation in (7).

The theorem can be rephrased by saying that the map ρ_c is injective on $S^\perp$ if $\mathcal{U}$ is not a cut, and the induced map is injective on the quotient vector space $S^\perp/\langle\omega^0\rangle$ if $\mathcal{U}$ is a cut. If $\mathcal{G}_\mathcal{U}$ is strongly connected, then the map ρ_c is an isomorphism between $S^\perp$ and $\widetilde{S}^\perp$ if $\mathcal{U}$ is not a cut, and between $S^\perp/\langle\omega^0\rangle$ and $\widetilde{S}^\perp$ if $\mathcal{U}$ is a cut.

Theorem 34 (i) gives that $\rho_c(S^\perp) \subseteq \widetilde{S}^\perp$, and by Theorem 34 (ii) there is equality if the graph $\mathcal{G}_\mathcal{U}$ is strongly connected. If $\mathcal{G}_\mathcal{U}$ is not strongly connected, then both $\widetilde{S}^\perp = \rho_c(S^\perp)$ and $\widetilde{S}^\perp \neq \rho_c(S^\perp)$ are possible. We present here some examples of these two situations.

Example 36. Consider the reaction network in Example 22. A minimal set of conservation laws is

$$T_1 = u_1 + u_2 + u_3 + u_4 \qquad T_3 = x_3 + x_4$$
$$T_2 = x_1 + x_2 + u_2 \qquad T_4 = x_1 + x_4 + u_2 + u_3 + 2u_4.$$

The map ρ_c gives rise to the following conservation laws for the reduced reaction network:

$$x_1 + x_2 = \widetilde{T}_2, \quad x_3 + x_4 = \widetilde{T}_3 \quad x_1 + x_4 = \widetilde{T}_4.$$

However, the reduced reaction network has no reactions and therefore $\widetilde{S}^\perp = \mathbb{R}^4$. This is an example of a non-strongly connected multidigraph $\mathcal{G}_\mathcal{U}$ where $\rho_c(S^\perp) \subsetneq \widetilde{S}^\perp$.

Example 37. Consider the reaction network

$$S_1 + U_1 \xrightarrow{\kappa_1(x,u)} S_2 + U_2 \qquad S_2 \xrightarrow{\kappa_2(x,u)} S_1$$

and the cut $\mathcal{U} = \{U_1, U_2\}$. A minimal set of conservation laws of the reaction network consists of $u_1 + u_2 = T_1$ and $x_1 + x_2 = T_2$. The multidigraph $\mathcal{G}_{\mathcal{U}}$ is

$$U_1 \xrightarrow{v_1(x)} U_2$$

which is acyclic. The reduced reaction network has one reaction $S_2 \longrightarrow S_1$, and a minimal set of conservation laws consists of $x_1 + x_2 = T_2$. This is an example of a non-strongly connected multidigraph $\mathcal{G}_{\mathcal{U}}$ where $\widetilde{S}^{\perp} = \rho_c(S^{\perp})$.

5.1 Iterative elimination

We conclude this section with a discussion on step-wise elimination.

Let $\emptyset = \mathcal{U}_0 \subset \mathcal{U}_1 \subset \mathcal{U}_2 \subset \cdots \subset \mathcal{U}_l = \mathcal{U}$ be increasing subsets of a noninteracting set $\mathcal{U}$. We want to compare the situation where $\mathcal{U}$ is eliminated at once (direct elimination) to the situation in which $\mathcal{U}$ is eliminated in steps by eliminating $\mathcal{U}_i \setminus \mathcal{U}_{i-1}$, $i = 1, \ldots, l$, iteratively (iterative elimination).

If each set $\mathcal{U}_i \setminus \mathcal{U}_{i-1}$, $i = 1, \ldots, l$, is a set of nodes of a connected component of $\mathcal{G}^*_{\mathcal{U}}$ (as defined in Section 4.3), then direct and iterative elimination yield the same result.

When $\mathcal{U}$ is not decomposable in this way, some care needs to be taken. If $\mathcal{U}$ is not a cut, then direct and iterative elimination still agree. When $\mathcal{U}$ is a cut, the rate functions depend on whether the reduced reaction network is obtained iteratively or not. We illustrate the two situations in Example 38.

Example 38. Consider the following reaction network with mass-action kinetics:

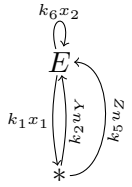
$$S_1 + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} Y \underset{k_4}{\overset{k_3}{\rightleftharpoons}} Z \xrightarrow{k_5} S_2 + E \xrightarrow{k_6} S_3 + E,$$

and the sets of noninteracting species:

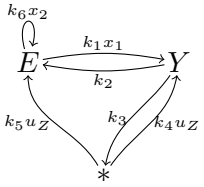
$$\mathcal{U}_1 = \{E\} \subset \mathcal{U}_2 = \{E, Y\} \subset \mathcal{U}_3 = \{E, Y, Z\},$$

among which only $\mathcal{U}_3$ is a cut. The multidigraphs $\mathcal{G}_{\mathcal{U}_i}$ are

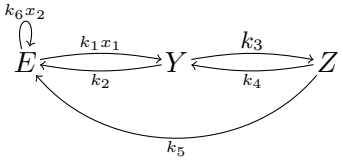
$G_{\mathcal{U}_1}$:



$G_{\mathcal{U}_2}$:



$G_{\mathcal{U}_3}$:



We denote by u_E, u_Y, u_Z the concentrations of E, Y, Z, respectively, and by x_i the concentration of S_i for $i = 1, 2, 3$. The reduced reaction network obtained by elimination of $\mathcal{U}_1 = \{E\}$ is

$$S_1 + Z \xrightarrow{\widetilde{\kappa}_1(x,u_Y,u_Z)} Y + S_2 \qquad Y \underset{\widetilde{\kappa}_3(x,u_Y,u_Z)}{\overset{\widetilde{\kappa}_2(x,u_Y,u_Z)}{\rightleftharpoons}} Z \qquad S_2 \xrightarrow{\widetilde{\kappa}_4(x,u_Y,u_Z)} S_3$$

with rate functions

$$\widetilde{\kappa}_1(x,u_Y,u_Z)=\frac{k_1k_5x_1u_Z}{k_1x_1}=k_5u_Z,\qquad \widetilde{\kappa}_2(x,u_Y,u_Z)=k_3u_Y,$$
$$\widetilde{\kappa}_3(x,u_Y,u_Z)=k_4u_Z,\qquad \widetilde{\kappa}_4(x,u_Y,u_Z)=\frac{k_6x_2(k_2u_Y+k_5u_Z)}{k_1x_1}.$$

In this case $\widetilde{\Omega}=\mathbb{R}^5_{\geq 0}\setminus\{x\in\mathbb{R}^5_{\geq 0}\mid x_1=0\}$, while $\Omega=\mathbb{R}^5_{\geq 0}$. Since $x_1\neq 0$ for all $x\in\widetilde{\Omega}$ we might cancel x_1 from $\widetilde{\kappa}_1(x,u_Y,u_Z)$.

The kinetics is $\mathcal{U}_2'$-linear with $\mathcal{U}_2'=\{Y\}$. From the first reduced reaction network we obtain a new reduced reaction network by elimination of $\mathcal{U}_2'$:

$$S_1+Z\xrightarrow{\widetilde{\kappa}_1'(x,u_Z)}S_2+Z\qquad S_2\xrightarrow{\widetilde{\kappa}_2'(x,u_Z)}S_3 \tag{25}$$

with rate functions

$$\widetilde{\kappa}_1'(x,u_z)=\frac{k_3k_5u_Z}{k_3}=k_5u_Z,\qquad \widetilde{\kappa}_2'(x,u_Z)=\frac{(k_2k_4+k_2k_5+k_3k_5)k_6u_Zx_2}{k_1k_3x_1}.$$

This is also the network we obtain by elimination of $\mathcal{U}_2$ directly.

If we eliminate the cut $\mathcal{U}_3'=\{Z\}$ from (25) with conservation law $u_Z=T$ or the cut $\mathcal{U}_3$ from the original reaction network with conservation law $u_E+u_Y+u_Z=T$, then we obtain the network

$$S_1\xrightarrow{\widetilde{\kappa}_1''(x)}S_2\xrightarrow{\widetilde{\kappa}_2''(x)}S_3.$$

In both cases the rate functions are obtained from $\widetilde{\kappa}_1'(x),\widetilde{\kappa}_2'(x)$: in the first case (iterative elimination) by letting $u_Z=T$, while in the second case (direct elimination) by letting

$$u_Z=\frac{k_1k_3x_1}{k_1(k_3+k_4+k_5)x_1+k_2k_4+k_2k_5+k_3k_5}.$$

Therefore, the kinetics of the reduced reaction network depends on whether the elimination is performed iteratively or not.

When $\mathcal{U}$ is not a cut and $\mathcal{G}_{\mathcal{U}}$ is connected, elimination of u only involves the equations $\dot{u}=0$ and no conservation laws. Therefore, iterative elimination is equivalent to solving a system of linear equations iteratively, and hence iterative and direct elimination coincide. When $\mathcal{U}$ is a cut this is not the case.

Let $\mathcal{U}$ be a cut such that $\mathcal{G}_{\mathcal{U}}$ satisfies (ST). Assume, for reasons of simplification, that $\mathcal{G}_{\mathcal{U}}$ is strongly connected. Further, assume that $\mathcal{U}_1=\{U_1,\dots,U_{m-1}\}$ and $\mathcal{U}_2=\mathcal{U}$. Theorem 17 holds for $\mathcal{U}_1$. Let $\varphi^1(x,u_m)$ be as in (11) for $\mathcal{U}_1$. Since $\mathcal{U}_1$ is not a cut, direct and iterative elimination yield the same result, in particular $\varphi^1(x,u_m)$ is not affected by the chosen procedure.

The system of equations for elimination of $\mathcal{U}$ is

$$\dot{u}_1=\cdots=\dot{u}_{m-1}=0$$
$$\sum_{i=1}^{m}u_i=T.$$

Using the first $m-1$ equations yields $u_i=\varphi_i^1(x,u_m)$ for $i=1,\dots,m-1$. Hence, the last equation translates into

$$\sum_{i=1}^{m-1}\varphi_i^1(x,u_m)+u_m=T. \tag{26}$$

If $\mathcal{U}_1$ is eliminated first, we obtain $u_i = \varphi_i^1(x, u_m)$ and there is only one equation left for elimination of $\mathcal{U} \setminus \mathcal{U}_1 = \{U_m\}$:

$$u_m = \widetilde{T}, \tag{27}$$

which follows from Theorem 34. By comparing (26) and (27), we conclude that the two procedures give different solutions for u_m. Hence, the kinetics of the reduced reaction network depends on the chosen procedure.

However, the reduced reaction network itself (without the kinetics) is independent of whether elimination is performed iteratively or not. To see this, consider $\mathcal{U}_1 \subset \mathcal{U}_2 = \mathcal{U}$ as above. The difference in the results of the two procedures is due to the conservation laws. Since $\mathcal{U}_1$ is not a cut, $*$ is a node of $\mathcal{G}_{\mathcal{U}_1}$. Since $\mathcal{U}_1 \cup \{U_m\}$ is a cut, any reaction $r \in \mathcal{R}$ such that $*$ is the source or target of $e(r)$ in $\mathcal{G}_{\mathcal{U}_1}$ involves U_m as a reactant or as a product. Hence, for each cycle in $\mathcal{G}_{\mathcal{U}_1}$ there is a cycle in $\mathcal{G}_{\mathcal{U}}$ with the same nodes, identifying $*$ with U_m, and whose edges correspond to the same reactions in $\mathcal{R}$. Besides, there may be cycles in $\mathcal{G}_{\mathcal{U}}$ corresponding to self-edges for U_m.

Let $\mathcal{G}_{\mathcal{U}\setminus\mathcal{U}_1}$ be the graph corresponding to the elimination of $\mathcal{U}\setminus\mathcal{U}_1 = \{U_m\}$ from the reduced reaction network $(\widetilde{\mathcal{C}}^1, \widetilde{\mathcal{R}}_1^1 \cup \widetilde{\mathcal{R}}_2^1)$ obtained by elimination of $\mathcal{U}_1$. The graph $\mathcal{G}_{\mathcal{U}\setminus\mathcal{U}_1}$ has a single node U_m and the edges are self-edges. There are two kinds of edges in $\mathcal{G}_{\mathcal{U}\setminus\mathcal{U}_1}$:

(1) Edges corresponding to reactions in $\widetilde{\mathcal{R}}_1^1$, that are defined by reactions in $\mathcal{R}$ with U_m involved in the reactant and the product, and corresponding to self-edges for U_m in $\mathcal{G}_{\mathcal{U}}$.

(2) Edges corresponding to reactions in $\widetilde{\mathcal{R}}_2^1$, that are defined by cycles in $\mathcal{G}_{\mathcal{U}_1}$ that contain $*$. These cycles correspond to cycles in $\mathcal{G}_{\mathcal{U}}$ that contain U_m and are not self-edges.

Let $(\widetilde{\mathcal{C}}^2, \widetilde{\mathcal{R}}_1^2 \cup \widetilde{\mathcal{R}}_2^2)$ be the reduced reaction network obtained by elimination of U_m from $(\widetilde{\mathcal{C}}^1, \widetilde{\mathcal{R}}_1^1 \cup \widetilde{\mathcal{R}}_2^1)$ (iterative elimination) and $(\widetilde{\mathcal{C}}, \widetilde{\mathcal{R}}_1 \cup \widetilde{\mathcal{R}}_2)$ the reduced reaction network obtained by elimination of $\mathcal{U}$ (direct elimination). A reaction $\widetilde{r} \in \widetilde{\mathcal{R}}_1$ belongs also to $\widetilde{\mathcal{R}}_1^2$, with possibly different rate function, since it is defined by a reaction in $\mathcal{R}_{\mathcal{U}}^{\perp}$. Let $\widetilde{r}_\sigma \in \widetilde{\mathcal{R}}_2$ be the reaction defined by the cycle σ. If σ contains U_m, then $\widetilde{r}_\sigma \in \widetilde{\mathcal{R}}_2^2$. If σ does not contain U_m, then $\widetilde{r}_\sigma \in \widetilde{\mathcal{R}}_2^1$ where $\widetilde{r}_\sigma$ is defined by the cycle σ in $\mathcal{G}_{\mathcal{U}_1}$ and thus $\widetilde{r}_\sigma \in \widetilde{\mathcal{R}}_1^2$.

Hence, $\widetilde{\mathcal{R}}_1 \cup \widetilde{\mathcal{R}}_2 \subseteq \widetilde{\mathcal{R}}_1^2 \cup \widetilde{\mathcal{R}}_2^2$. The other inclusion is deduced by similar arguments.

6 Examples

We conclude by discussing some biological examples. In this section we abuse notation and use x for the vector of concentrations of the original set of species as well as of the species in the reduced reaction network.

6.1 Mechanism with two substrates, two products and one enzyme

We consider the following mechanism discussed in [3, Chapter 5]. The mechanism consists of two substrates that are converted into two products through a series of reactions. It is a particular case of a *bi-bi* model in the notation introduced by Cleland [9]:

$$\begin{array}{lll}
\textit{Substrate binding:} & E + A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} EA & EA + B \underset{k_4}{\overset{k_3}{\rightleftharpoons}} EAB \\
\textit{Transformation:} & EAB \underset{k_6}{\overset{k_5}{\rightleftharpoons}} EPQ & \\
\textit{Product release:} & EPQ \underset{k_8}{\overset{k_7}{\rightleftharpoons}} EQ + P & EQ \underset{k_{10}}{\overset{k_9}{\rightleftharpoons}} E + Q
\end{array}$$

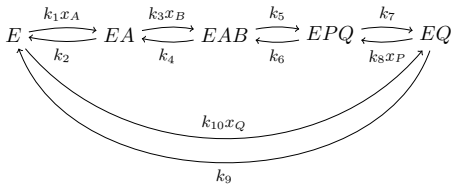
We assume mass-action kinetics. A minimal set of conservation laws is

$$\begin{aligned}
T_E &= x_E + x_{EA} + x_{EAB} + x_{EPQ} + x_{EQ}\\
T_{A+Q} &= x_A + x_Q + x_{EA} + x_{EAB} + x_{EPQ} + x_{EQ}\\
T_{B+Q} &= x_B + x_Q + x_{EAB} + x_{EPQ} + x_{EQ}\\
T_{B+P} &= x_B + x_P + x_{EAB} + x_{EPQ}.
\end{aligned}$$

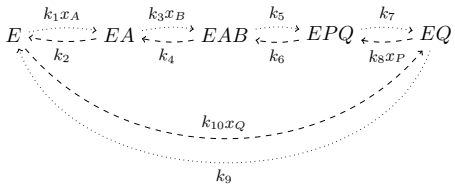
We consider the reduced reaction network obtained by elimination of the cut

$$\mathcal{U} = \{E, EA, EAB, EPQ, EQ\}.$$

The multidigraph $\mathcal{G}_{\mathcal{U}}$ is



Only two of the cycles in $\mathcal{G}_{\mathcal{U}}$ define reactions for the reduced reaction network:



We apply the procedure for the dotted cycle:

$$\begin{array}{rcl}
\cancel{E} + A & \longrightarrow & \cancel{EA}\\
\cancel{EA} + B & \longrightarrow & \cancel{EAB}\\
\cancel{EAB} & \longrightarrow & \cancel{EPQ}\\
\cancel{EPQ} & \longrightarrow & \cancel{EQ} + P\\
\cancel{EQ} & \longrightarrow & \cancel{E} + Q\\
\hline
A + B & \xrightarrow{\widetilde{\kappa}_1(x)} & P + Q
\end{array}$$

the second cycle is just the one given by the reverse reactions.

Note that for both cycles, $\Gamma(\sigma) = \{\sigma\}$, see Remark 21. Therefore, the reduced reaction network is:

$$A + B \underset{\widetilde{\kappa}_2(x)}{\overset{\widetilde{\kappa}_1(x)}{\rightleftharpoons}} P + Q$$

where

$$\widetilde{\kappa}_1(x) = q(x)k_1k_3k_5k_7k_9x_Ax_B \qquad \widetilde{\kappa}_2(x) = q(x)k_2k_4k_6k_8k_{10}x_Px_Q$$

and

$$\begin{aligned}
q(x) =& T_E\Big((k_4k_6 + k_4k_7 + k_5k_7)k_2k_9 + k_2k_4k_6k_8x_P + k_3k_5k_7k_9x_B\\
&+ (k_4k_6 + k_4k_7 + k_5k_7)k_2k_{10}x_Q + (k_4k_6 + k_5k_7 + k_4k_7)k_1k_9x_A\\
&+ (k_5k_9 + k_7k_9 + k_5k_7 + k_6k_9)k_1k_3x_Ax_B + k_3k_5k_7k_{10}x_Bx_Q\\
&+ k_1k_4k_6k_8x_Ax_P + (k_5k_2 + k_4k_2 + k_6k_2 + k_6k_4)k_8k_{10}x_Px_Q\\
&+ (k_5 + k_6)k_1k_3k_8x_Ax_Bx_P + (k_5 + k_6)k_3k_8k_{10}x_Bx_Px_Q\Big)^{-1}.
\end{aligned}$$

In this case $\widetilde{\Omega} = \mathbb{R}^4_{\geq 0}$. A minimal set of conservation laws for the reduced reaction network is

$$x_A + x_Q = \widetilde{T}_{A+Q}, \qquad x_B + x_Q = \widetilde{T}_{B+Q}, \qquad x_B + x_P = \widetilde{T}_{B+P}.$$

These conservation laws are obtained by projecting the conservation laws of the original reaction network into the coordinates of the non-eliminated species.

Note that the reduced reaction network has the same reactions as the reduced reaction network for the *ping-pong bi-bi* mechanism considered in the introduction. The difference between the two reduced reaction networks lays in the terms present in the factor $q(x)$.

6.2 Intermediates

Our construction generalises the reduction of intermediates in [6]. In [6], an intermediate is defined as a species Y in a reaction network that is created and dissociated in isolation, that is, it is produced in at least one reaction, consumed in at least one reaction and is not involved in any complex other than Y. A set of intermediates, $\mathcal{Y}$, is a subset of the species set and at the same time a subset of the set of complexes (under the identification of complexes with linear combinations of species). It is moreover assumed that every intermediate belongs to a path between two complexes in $\mathcal{C} \setminus \mathcal{Y}$.

Any set of intermediates is a noninteracting set that is not a cut and assumption (ST) holds on $\mathcal{G}_{\mathcal{Y}}$. By assuming mass-action kinetics, it is possible to eliminate $\mathcal{Y}$ using the results in Theorem 17.

We make some observations on the multidigraph $\mathcal{G}_{\mathcal{Y}}$:

(i) If $r(e)$ involves some species in $\mathcal{Y}^c$, then $s(e) = *$ or $t(e) = *$.

(ii) If the label of an edge e depends on some species concentration (i.e. is not constant), then $s(e) = *$.

(iii) If τ is a spanning tree rooted at $*$, then $\pi(\tau)$ is a product of reaction rate constants k_i.

(iv) The edge e with $s(e) = *$ of a spanning tree rooted at $Y \in \mathcal{Y}$ is the only edge of the tree whose label depends on some species concentration.

By (i), any cycle $\sigma \in \widetilde{\Delta}$ must contain node $*$. The reactant of $\widetilde{r}_\sigma$ is then the reactant of the reaction corresponding to the edge with source $*$, and the product of $\widetilde{r}_\sigma$ is the product of the reaction corresponding to the edge with target $*$.

By (iii) and (iv),

$$q(x) = g(k)^{-1}, \qquad \text{and} \qquad \Pi(\sigma) = x^{y_{\widetilde{r}_\sigma}} f(k),$$

where $f(k), g(k)$ are polynomials in the reaction rate constants, such that $q(x)$ is constant in x. Therefore, the reduced reaction network obtained by elimination of $\mathcal{Y}$ is a reaction network with mass-action kinetics. There is a reaction between two complexes in $\mathcal{C} \setminus \mathcal{Y}$ in the reduced reaction network if and only if the reaction is already in the original network or there is a path between them through intermediates.

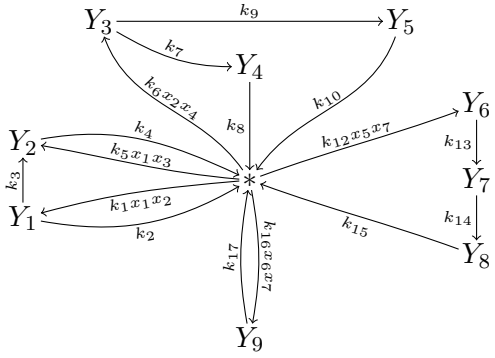
Example 39. Consider the following reaction network with mass-action kinetics:

$$S_1 + S_2 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} Y_1 \xrightarrow{k_3} Y_2 \underset{k_5}{\overset{k_4}{\rightleftharpoons}} S_1 + S_3 \qquad Y_3 \xrightarrow{k_9} Y_5 \xrightarrow{k_{10}} S_4 + S_6$$

$$S_2 + S_4 \xrightarrow{k_6} Y_3 \xrightarrow{k_7} Y_4 \xrightarrow{k_8} S_4 + S_5 \qquad S_3 + S_7 \xrightarrow{k_{11}} S_2 + S_7$$

$$S_5 + S_7 \xrightarrow{k_{12}} Y_6 \xrightarrow{k_{13}} Y_7 \xrightarrow{k_{14}} Y_8 \xrightarrow{k_{15}} S_2 + S_7 \qquad S_6 + S_7 \xrightarrow{k_{16}} Y_9 \xrightarrow{k_{17}} S_2 + S_7.$$

A set of intermediates in this network is $\mathcal{Y} = \{Y_1, Y_2, \ldots, Y_9\}$. The corresponding multidigraph $\mathcal{G}_\mathcal{Y}$ is



By Proposition 30, the reduced reaction network obtained by elimination of $\mathcal{Y}$ can be computed by iteratively eliminating the sets of nodes $\mathcal{Y}_1 = \{Y_1, Y_2\}$, $\mathcal{Y}_2 = \{Y_3, Y_4, Y_5\}$, $\mathcal{Y}_3 = \{Y_6, Y_7, Y_8\}$ and $\mathcal{Y}_4 = \{Y_9\}$. The multidigraphs $\mathcal{G}_{\mathcal{Y}_1}$, $\mathcal{G}_{\mathcal{Y}_2}$, $\mathcal{G}_{\mathcal{Y}_3}$ and $\mathcal{G}_{\mathcal{Y}_4}$ correspond to the left-subgraph, top-subgraph, right-subgraph and bottom-subgraph of $\mathcal{G}_\mathcal{Y}$, respectively. We find the cycles in each of the multidigraphs and obtain the following reduced reaction network with the specified rate functions:

$$S_1 + S_2 \xrightarrow{\widetilde{\kappa}_1(x)} S_1 + S_3 \qquad S_4 + S_2 \xrightarrow{\widetilde{\kappa}_2(x)} S_4 + S_5 \qquad S_2 + S_4 \xrightarrow{\widetilde{\kappa}_3(x)} S_4 + S_6$$
$$S_3 + S_7 \xrightarrow{\widetilde{\kappa}_4(x)} S_2 + S_7 \qquad S_7 + S_5 \xrightarrow{\widetilde{\kappa}_5(x)} S_2 + S_7 \qquad S_6 + S_7 \xrightarrow{\widetilde{\kappa}_6(x)} S_2 + S_7$$

where

$$\widetilde{\kappa}_1(x) = \frac{k_1k_3k_4x_1x_2}{k_4(k_2+k_3)} = \frac{k_1k_3x_1x_2}{(k_2+k_3)}, \qquad \widetilde{\kappa}_4(x) = k_{11}x_3x_7,$$
$$\widetilde{\kappa}_2(x) = \frac{k_6k_9k_{10}x_5x_7}{k_8k_{10}(k_7+k_9)} = \frac{k_6k_9x_5x_7}{k_8(k_7+k_9)}, \qquad \widetilde{\kappa}_5(x) = \frac{k_{12}k_{13}k_{14}k_{15}x_5x_7}{k_{13}k_{14}k_{15}} = k_{12}x_5x_7,$$
$$\widetilde{\kappa}_3(x) = \frac{k_6k_7k_8x_5x_7}{k_8k_{10}(k_7+k_9)} = \frac{k_6k_7x_5x_7}{k_{10}(k_7+k_9)}, \qquad \widetilde{\kappa}_6(x) = \frac{k_{16}k_{17}x_6x_7}{k_{17}} = k_{16}x_6x_7.$$

Indeed, the reduced reaction network has mass-action kinetics as well.

6.3 Post-translational modification networks

A common feature of signaling systems is the incorporation of Post-Translational Modifications (PTMs). A PTM consists of the attachment of some chemical group to a protein, after it has been translated. The most common example is phosphorylation, as will be discussed in Example 42.

PTMs are used to activate or inactivate proteins in order to propagate the signal through the system. PTM networks, that is, a network combining several PTMs, have attracted considerable attention due to their abundance in nature [11] and regular form [18].

In [18, 7] these networks are considered from the point of view of variable elimination. In these works, substrates and intermediates are eliminated from the system of steady state equations and the conservation laws for the substrate, providing a system of equations depending on the enzyme concentrations only.

Here we study the reduced reaction network obtained by elimination of the sets of enzymes and intermediates, and obtain a reaction network that involves only transformations among the substrates.

We start by giving the definition of a PTM network, which differs slightly from the definition in [18, 7]. The species set of a PTM network is the disjoint union of three non-empty sets of species:

- a set of substrates $\mathcal{S} = \{S_1, \ldots, S_p\}$,
- a set of enzymes $\mathcal{E} = \{E_1, \ldots, E_{m_1}\}$ and
- a set of intermediates $\mathcal{Y} = \{Y_1, \ldots, Y_{m_2}\}$, in the sense of Subsection 6.2.

The set of reactions consists of reactions of one of these five types:

- $S_i + E_j \longrightarrow Y_l$,
- $Y_l \longrightarrow S_i + E_j$,
- $Y_i \longrightarrow Y_j$
- $S_i \longrightarrow S_j$ and
- $S_i + E_j \longrightarrow S_l + E_j$.

We assume that any path

$$S_{i_1} + E_{j_1} \longrightarrow Y_{l_1} \longrightarrow \cdots \longrightarrow Y_{l_t} \longrightarrow S_{i_2} + E_{j_2} \tag{28}$$

in the reaction network, for which only the end complexes are non-intermediates, satisfies $j_1 = j_2$. This provides a decomposition of the set of intermediates $\mathcal{Y}$ into m_1 disjoint subsets (some of these subsets may be empty) according to the enzyme that ultimately forms them (or to which they dissociate). We allow for reactions like $S_1 + E \to S_2 + E$ or $S_1 \to S_2$, that are not considered in [18, 7].

We observe that the set of enzymes and intermediates, $\mathcal{U} = \mathcal{E} \cup \mathcal{Y}$, is a cut. The multidigraph $\mathcal{G}_{\mathcal{U}}$ has a connected component for each enzyme. By the hypothesis on the set of intermediates, each connected component is strongly connected.

The next proposition describes the reaction of the reduced reaction network obtained by elimination of $\mathcal{U}$. Recall the decomposition $\widetilde{\mathcal{R}} = \widetilde{\mathcal{R}}_1 \cup \widetilde{\mathcal{R}}_2$ of the set of reactions of the reduced reaction network, given in Definition 23.

A reaction $S_i \to S_j$ in the reduced reaction network belongs to $\widetilde{\mathcal{R}}_1$ if it already belongs to $\mathcal{R}$. Therefore, in the following proposition we only consider $\widetilde{\mathcal{R}}_2$.

Proposition 40. *Let $(\mathcal{C}, \mathcal{R})$ be a PTM network and $(\widetilde{\mathcal{C}}, \widetilde{\mathcal{R}})$ be the reduced reaction network obtained by elimination of $\mathcal{U} = \mathcal{E} \cup \mathcal{Y}$. Let $S_{i_1}, S_{i_2} \in \mathcal{S}$ be two substrates. The reaction $S_{i_1} \to S_{i_2}$ belongs to $\widetilde{\mathcal{R}}_2$ if and only if there is a path as* (28) *from $S_{i_1} + E$ to $S_{i_2} + E$ in $(\mathcal{C}, \mathcal{R})$.*

Proof. The only edges in $\mathcal{G}_{\mathcal{U}}$ whose associated reaction involves a substrate (that is, a species in $\mathcal{U}^c$) are those with E as source or target. Recall that the reduced reaction network when $\mathcal{G}_{\mathcal{U}}$ is not connected is obtained by considering each connected component of $\mathcal{G}_{\mathcal{U}}$ independently (cf. Subsection 4.3).

$\Leftarrow$) Consider one connected component of $\mathcal{G}_{\mathcal{U}}$ with set of nodes $\mathcal{U}'$. Then $\mathcal{U}'$ is a cut and contains a unique enzyme $E \in \mathcal{E}$. Let

$$S_{i_1} + E \longrightarrow Y_{j_1} \longrightarrow \ldots \longrightarrow Y_{j_s} \longrightarrow E + S_{i_2}, \qquad s \geq 0,$$

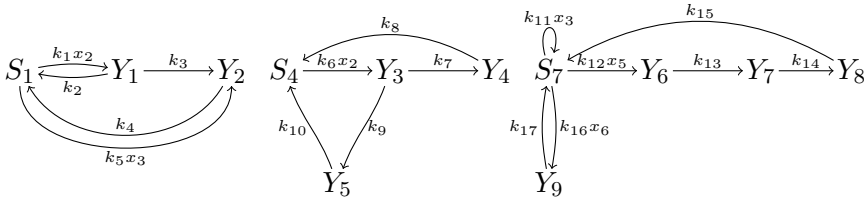
be a path in $(\mathcal{C}, \mathcal{R})$. This path defines a cycle σ in $\mathcal{G}_{\mathcal{U}'}$ with nodes E, $Y_{j_1}, \ldots, Y_{j_s}$ (it is a self-edge if $s = 0$). This cycle is such that $\Gamma(\sigma) \neq \emptyset$ (by Proposition 20) and gives rise to a reaction

$$S_{i_1} \longrightarrow S_{i_2} \in \widetilde{\mathcal{R}}_2.$$

$\Rightarrow$) If $S_{i_1} \longrightarrow S_{i_2} \in \widetilde{\mathcal{R}}_2$, then there is a cycle $\sigma \in \widetilde{\Delta}$ defining it. The cycle σ must contain a unique enzyme $E \in \mathcal{E}$. It follows that one of the edges in σ corresponds to a reaction with reactant $S_{i_1} + E$, and one of edges corresponds to a reaction with product $S_{i_2} + E$. All other edges correspond to reactions between intermediates. The reactions corresponding to the edges in the cycle give the claimed path from $S_{i_1} + E$ to $S_{i_2} + E$. □

We conclude that if two substrates do not interact with a common enzyme E, there is no reaction between them in $\widetilde{\mathcal{R}}_2$.

Example 41. We consider again the network in Example 39. With $\mathcal{Y}$ as in the example, $\mathcal{E} = \{S_1, S_4, S_7\}$ a set of enzymes and $\mathcal{S} = \{S_2, S_3, S_5, S_6\}$ a set of substrates, the network is a PTM network. Let $\mathcal{U} = \mathcal{E} \cup \mathcal{Y}$. The multidigraph $\mathcal{G}_{\mathcal{U}}$ in this case is



By considering the three components of $\mathcal{G}_{\mathcal{U}}$ separately, we obtain the following reduced reaction network by elimination of $\mathcal{U}$:

$$S_2 \rightleftharpoons S_3 \qquad S_2 \rightleftharpoons S_5 \qquad S_2 \rightleftharpoons S_6$$

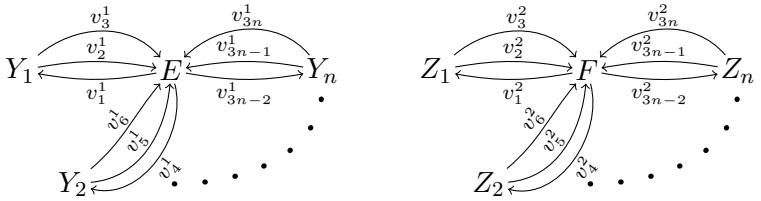
This reaction network is in agreement with the statement of Proposition 40.

Example 42 (n-site phosphorylation system)**.** Phosphorylation and dephosphorylation of proteins, in one or multiple sites, play an important role in many intracellular processes, such as signal transduction, cell-cycle control and nuclear signal integration. Mathematical models of such mechanisms have often been studied, see for instance [14, 18, 8].

We consider here an n-site sequential distributive phosphorylation mechanism, which consists of a substrate S that contains n ordered phosphorylation sites. We let S_0 denote the unphosphorylated form and S_i denote the phosphorylated form in which sites 1 to i are phosphorylated. We assume there is a kinase E that catalyses all phosphorylation steps and, similarly, a phosphatase F that catalyses dephosphorylation steps. The reaction network associated with this system is

$$S_0 + E \underset{\kappa_2^1(x)}{\overset{\kappa_1^1(x)}{\rightleftharpoons}} Y_1 \xrightarrow{\kappa_3^1(x)} S_1 + E \underset{\kappa_5^1(x)}{\overset{\kappa_4^1(x)}{\rightleftharpoons}} Y_2 \longrightarrow \ldots \rightleftharpoons Y_n \xrightarrow{\kappa_{3n}^1(x)} S_n + E$$
$$S_n + F \underset{\kappa_2^2(x)}{\overset{\kappa_1^2(x)}{\rightleftharpoons}} Z_1 \xrightarrow{\kappa_3^2(x)} S_{n-1} + F \underset{\kappa_5^2(x)}{\overset{\kappa_4^2(x)}{\rightleftharpoons}} Z_2 \longrightarrow \cdots \rightleftharpoons Z_n \xrightarrow{\kappa_{3n}^2(x)} S_0 + F.$$

For the cut $\mathcal{U} = \{E, F, Y_1, \ldots, Y_n, Z_1, \ldots, Z_n\}$ and a $\mathcal{U}$-linear kinetics, the multidigraph $\mathcal{G}_{\mathcal{U}}$ has two connected components:



where we have omitted the dependence of v_i on x to simplify the notation. Each connected component corresponds to a cut, with total amount T^1, T^2, respectively. The cycles in $\widetilde{\Delta}$ in $\mathcal{G}_{\mathcal{U}}$ are:

$$\sigma_i^1\colon E \underset{v_{3i}^1}{\overset{v_{3i-1}^1}{\rightleftarrows}} Y_i \qquad\qquad \sigma_i^2\colon F \underset{v_{3i}^2}{\overset{v_{3i-1}^2}{\rightleftarrows}} Z_i$$

for $i = 1, \ldots, n$. By Proposition 40, the reduced reaction network obtained by elimination of $\mathcal{U}$ is:

$$S_0 \underset{\widetilde{\kappa}_n^2(x)}{\overset{\widetilde{\kappa}_1^1(x)}{\rightleftarrows}} S_1 \underset{\widetilde{\kappa}_{n-1}^2(x)}{\overset{\widetilde{\kappa}_2^1(x)}{\rightleftarrows}} \cdots \underset{\widetilde{\kappa}_1^2(x)}{\overset{\widetilde{\kappa}_n^1(x)}{\rightleftarrows}} S_n.$$

For $l = 1, 2$ and $i = 1, \ldots, n$ the reaction with rate function $\widetilde{\kappa}_i^l(x)$ corresponds to the cycle σ_i^l and is as follows:

$$\widetilde{\kappa}_i^l(x) = q_l(x) v_{3i}^l v_{3i-2}^l \prod_{j=1, j\neq i}^{n} (v_{3j-1}^l + v_{3j}^l)$$

$$q_l(x) = T^l \Big(\prod_{j=1}^{n} (v_{3j-1}^l + v_{3j}^l) + \sum_{i=1}^{n} v_{3i-2}^l \prod_{j=1, j\neq i}^{n} (v_{3j-1}^l + v_{3j}^l) \Big)^{-1}.$$

6.4 Relation to previous work

Similar ideas for the elimination of species and subsequent simplification of reaction networks have been proposed in the literature. We will focus on the techniques proposed in [15] and explained also in [16] and [17]. We will discuss similarities and differences following the exposition in [16].

In [16] the authors outline some applications of graph theory to the theory of reaction networks. We note that the reversible reactions are treated as one reaction in [16] while we treat them as two separate reactions in the present paper. First, they introduce a graph called the *kinetic graph* (proposed in [15]) whose nodes represent the so called *intermediates*. The kinetic graph coincides with the graph $\mathcal{G}_{\mathcal{U}}$ introduced in Definition 14 (up to the treatment of reversible reaction as two reactions or one reaction) and the so called intermediates form a set of noninteracting species in our terminology. Hence, the main tool for the graphic application is the same graph used in the present paper. That (ST) holds and the kinetics is $\mathcal{U}$-linear is also required for the algorithms in [16] to apply.

In [16] the authors describe the elimination of intermediates from a reaction network in order to find a minimal mechanism that allows the computation of the production rates. The differences of their procedure and our work arise from the details in the treatment of the reactions and the rates.

First, in [16] not only one reduced reaction network is obtained, as it is the case in the present paper but an infinite number of them. Any linear combination of the original reactions that cancel the intermediates is a possible reaction in a reduced reaction network. The

set of such reactions defines a vector subspace of S, of which a basis is chosen. Therefore, the reactions in a reduced reaction network are independent as vectors in the stoichiometric subspace, and their number is minimal.

Second, the conditions imposed for the computation of the rate functions are different. The goal in the present work has been to find a set of reactions with respective rate functions that fulfilled (12). In [16], the rate functions for the reactions in a chosen reduced reaction network are found by imposing the Horiuti-Temkin equation (see (29)), which involves the rate functions for the original and reduced reaction networks but does not relate to the stochiometry of the reactions. As a consequence of the Houriti-Temkin equation, the ODE system (12) is also satisfied for the reactions and rate functions given in [16]. We discuss next in further detail the differences on this particular point. Assume for the discussion below that $\mathcal{R}_{\mathcal{U}}^{\perp} = \emptyset$.

Let $\sigma \in \tilde{\Delta}$. Assume that the set of edges in $\mathcal{G}_{\mathcal{U}}$ is ordered. Let ν_σ be the vector with $(\nu_\sigma)_i = 1$ if e_i is an edge of the cycle and $(\nu_\sigma)_i = 0$ otherwise. Assume also that the set of cycles $\widetilde{\Delta} = \{\sigma_1, \ldots, \sigma_{|\widetilde{\Delta}|}\}$ is ordered, and let $\widetilde{H}$ be the $|\widetilde{\Delta}| \times \ell$ matrix whose ith row is ν_{σ_i}. With this notation, the Horiuti-Temkin equation for the reduced reaction network in Definition 23, considering reversible reactions as two irreversible reactions, reads

$$\kappa(x, \varphi(x)) = \widetilde{H}^t \widetilde{\kappa}(x). \tag{29}$$

Componentwise, this condition translates into our notation as

$$\kappa_i(x, \varphi(x)) = \sum_{(\nu_{\sigma_j})_i \neq 0} \widetilde{\kappa}_j(x) = q(x) \sum_{\sigma \in \widetilde{\Delta}, e(r_i) \in \sigma} \Pi(\sigma) \qquad i = 1, \ldots, \ell.$$

This condition is not necessarily satisfied for the rate functions in Definition 23.

We show this in the case $\mathcal{U}$ is a cut. Assume that (ST) holds and the kinetics is $\mathcal{U}$-linear. Let $U_j \in \mathcal{U}$ be involved in the reactant of $r_i \in \mathcal{R}$, for $i \in \{1, \ldots, \ell\}$. Then

$$\kappa_i(x, \varphi(x)) = \varphi_j(x) v_i(x) = q(x) \sum_{\tau \in \Theta(U_j)} \pi(\tau) v_i(x) = (\star).$$

The terms in the sum are the labels of elements in $\Gamma(\sigma)$ for cycles $\sigma \in \Delta$ that contain the edge $e(r_i)$. Using Proposition 20 it translates into

$$(\star) = q(x) \sum_{\sigma \in \Delta, \Gamma(\sigma) \neq \emptyset, e(r_i) \in \sigma} \Pi(\sigma) \geq q(x) \sum_{\sigma \in \widetilde{\Delta}, e(r_i) \in \sigma} \Pi(\sigma).$$

Since $\sum_{e \in \sigma} \rho_c(y'_{r(e)} - y_{r(e)}) \neq 0$ for $\sigma \in \widetilde{\Delta}$, the Horiuti-Temkin equation is not necessarily satisfied, because the sum in the left-hand side may involve more terms than the sum in the right-hand side.

Let A^c be the matrix given by the first p rows of the stoichiometric matrix A. The stoichiometric matrix of $(\widetilde{\mathcal{C}}, \widetilde{\mathcal{R}})$ is $\widetilde{A} = A^c \widetilde{H}^t$ by Definition 23. By equation (6) and Theorem 24, the kinetics defined in Definition 23 satisfy

$$\rho_c \left(A \kappa(x, \varphi(x)) \right) = A^c \widetilde{H}^t \widetilde{\kappa}(x).$$

This equation is similar to the Horiuti-Temkin equation once stochiometry is introduced.

Finally, the kinetics obtained in [16] is not necessarily standard if the kinetics of the original reaction network is, contrary to our kinetics (cf. Section 4.4). We illustrate this with one example.

We show in further detail the differences of the two approaches using the main example in [16].

Example 43. Consider the following reaction network with mass-action kinetics

$$S_1 + S_2 + S_6 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} S_3 + S_7 \qquad S_7 \underset{k_4}{\overset{k_3}{\rightleftharpoons}} S_4 + S_6$$
$$S_4 + S_7 \underset{k_6}{\overset{k_5}{\rightleftharpoons}} S_1 + S_5 + S_6$$

The correspondence with [16] is as follows:

$$S_1 = \mathrm{C} \qquad S_2 = \mathrm{H_2O} \qquad S_3 = \mathrm{H_2} \qquad S_4 = \mathrm{CO} \qquad S_5 = \mathrm{CO_2}$$
$$S_6 = \mathrm{Z_1} \qquad S_7 = \mathrm{COZ_1}.$$

The set of eliminated species is the cut $\mathcal{U} = \{S_6, S_7\}$. The reduced reaction network obtained by our procedure is

$$S_1 + S_2 \underset{\widetilde{\kappa}_2(x)}{\overset{\widetilde{\kappa}_1(x)}{\rightleftharpoons}} S_3 + S_4 \qquad S_1 + S_2 + S_4 \underset{\widetilde{\kappa}_4(x)}{\overset{\widetilde{\kappa}_3(x)}{\rightleftharpoons}} S_1 + S_3 + S_5$$
$$2S_4 \underset{\widetilde{\kappa}_6(x)}{\overset{\widetilde{\kappa}_5(x)}{\rightleftharpoons}} S_1 + S_5$$

with kinetics

$$\widetilde{\kappa}_1(x) = q(x)k_1k_3x_1x_2 \qquad \widetilde{\kappa}_2(x) = q(x)k_2k_4x_3x_4$$
$$\widetilde{\kappa}_3(x) = q(x)k_1k_5x_1x_2x_4 \qquad \widetilde{\kappa}_4(x) = q(x)k_2k_6x_1x_3x_5$$
$$\widetilde{\kappa}_5(x) = q(x)k_4k_5x_4^2 \qquad \widetilde{\kappa}_6(x) = q(x)k_3k_6x_1x_5$$

and

$$q(x) = \frac{T}{k_1x_1x_2 + k_6x_1x_5 + k_2x_3 + (k_4 + k_5)x_4 + k_3}.$$

The kinetics is standard for the reduced reaction network.

One of the reduced reaction networks obtained in [16] is

$$S_1 + S_2 \underset{\widetilde{\kappa}'_2(x)}{\overset{\widetilde{\kappa}'_1(x)}{\rightleftharpoons}} S_3 + S_4 \qquad S_2 + S_4 \underset{\widetilde{\kappa}'_4(x)}{\overset{\widetilde{\kappa}'_3(x)}{\rightleftharpoons}} S_3 + S_5, \tag{30}$$

with kinetics

$$\widetilde{\kappa}'_1(x) - \widetilde{\kappa}'_2(x) = q(x)(k_1k_3x_1x_2 - k_2k_4x_3x_4 + k_3k_6x_1x_5 - k_4k_5x_4^2)$$
$$\widetilde{\kappa}'_3(x) - \widetilde{\kappa}'_4(x) = q(x)(k_1k_5x_1x_2x_4 - k_2k_6x_1x_3x_5 - k_3k_6x_1x_5 + k_4k_5x_4^2)$$

and $q(x)$ as above.

In [16] the rates of the reactions are considered by pairs of reversible reactions, that is why we obtain $\widetilde{\kappa}'_1(x) - \widetilde{\kappa}'_2(x)$ by their algorithm. By collecting the terms according to its sign we deduce that

$$\widetilde{\kappa}'_1(x) = q(x)(k_1k_3x_1x_2 + k_3k_6x_1x_5) \qquad \widetilde{\kappa}'_2(x) = q(x)(k_2k_4x_3x_4 + k_4k_5x_4^2).$$

We observe that the kinetics is not standard for the reaction network (30).

Acknowledgements MS, EF, CW are supported by The Lundbeck Foundation (Denmark). EF and CW acknowledge funding from the Danish Research Council of Independent Research. EF has been supported by the project MTM2012-38122-C03-01/FEDER and MS by MTM2012-38122-C03-02/FEDER from the Ministerio de Economía y Competitividad, Spain.

References

[1] Berge, C.: Graphs. Amsterdam, NL: North-Holland (1985)

[2] Briggs, G., Haldane, J.: A note on the kinetics of enzyme action. Biochem J. **19** (1925)

[3] Cornish-Bowden, A.: Fundamentals of Enzyme Kinetics, third edn. Portland Press, London (2004)

[4] Feinberg, M.: Lectures on chemical reaction networks. Available online at `http://www.crnt.osu.edu/LecturesOnReactionNetworks` (1980)

[5] Feliu, E., Wiuf, C.: Variable elimination in chemical reaction networks with mass-action kinetics. SIAM J. Appl. Math. **72**, 959–981 (2012)

[6] Feliu, E., Wiuf, C.: Simplifying biochemical models with intermediate species. J. R. S. Interface **10**, 20130,484 (2013)

[7] Feliu, E., Wiuf, C.: Variable elimination in post-translational modification reaction networks with mass-action kinetics. J. Math. Biol. **66**(1), 281–310 (2013)

[8] Flockerzi, D., Holstein, K., Conradi, C.: N-site Phosphorylation Systems with 2N-1 Steady States. Bulletin of Mathematical Biology **76**(8), 1892–1916 (2014). DOI 10.1007/s11538-014-9984-0. URL `http://dx.doi.org/10.1007/s11538-014-9984-0`

[9] Frey, P., Hegeman, A.: Enzymatic Reaction Mechanisms. Oxford University Press (2007)

[10] Gunawardena, J.: A linear framework for time-scale separation in nonlinear biochemical systems. PLoS ONE **7**(5), e36,321 (2012)

[11] Huang, C.Y., Ferrell, J.E.: Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc. Natl. Acad. Sci. U.S.A. **93**, 10,078–10,083 (1996)

[12] King, E.L., Altman, C.: A schematic method of deriving the rate laws for enzyme-catalyzed reactions. J. Phys. Chem. **60**, 1375–1378 (1956)

[13] Radulescu, O., Gorban, A.N., Zinovyev, A., Lilienbaum, A.: Robust simplifications of multiscale biochemical networks. BMC Syst Biol **2**, 86 (2008)

[14] Salazar, C., Hofer, T.: Multisite protein phosphorylation–from molecular mechanisms to kinetic models. FEBS J. **276**(12), 3177–3198 (2009)

[15] Temkin, M.: Graphical method for the derivation of the rate laws of complex reactions. Dokl. Akad. Nauk SSSR **165** (1965)

[16] Temkin, O.N., Bonchev, D.G.: Application of graph theory to chemical kinetics: Part 1. kinetics of complex reactions. Journal of Chemical Education **69**(7), 544 (1992). DOI 10.1021/ed069p544. URL `http://dx.doi.org/10.1021/ed069p544`

[17] Temkin, O.N., Zeigarnik, A.V., Bonchev, D.: Chemical Reaction Networks: A Graph-Theoretical Approach. CRC Press (1996)

[18] Thomson, M., Gunawardena, J.: The rational parameterization theorem for multisite post-translational modification systems. J. Theor. Biol. **261**, 626–636 (2009)

[19] Walther, H.: Ten Applications of Graph Theory. Springer (1985)

[20] Wong, J.T., Hanes, C.S.: Kinetic formulations for enzymic reactions involving two substrates. Can J Biochem Physiol **40**, 763–804 (1962)