Realizations of kinetic differential equations

Gheorghe Craciun *   Matthew D. Johnston †   Gábor Szederkényi ‡

Elisa Tonello §   János Tóth ¶   Polly Y. Yu ‖

Abstract

The induced kinetic differential equations of a reaction network endowed with mass action type kinetics is a system of polynomial differential equations. The problem studied here is: Given a system of polynomial differential equations, is it possible to find a network which induces these equations; in other words: is it possible to find a *kinetic realization* of this system of differential equations? If yes, can we find a network with some chemically relevant properties (implying also important dynamic consequences), such as reversibility, weak reversibility, zero deficiency, detailed balancing, complex balancing, mass conservation, etc.? The constructive answers presented to a series of questions of the above type are useful when fitting differential equations to datasets, or when trying to find out the dynamic behavior of the solutions of differential equations. It turns out that some of these results can be applied when trying to solve seemingly unrelated mathematical problems, like the existence of positive solutions to algebraic equations.

Keywords: kinetic equations, reversibility, weak reversibility, mass action kinetics, reaction networks, realizations

1 Introduction

Our goal here is to find reaction networks inducing a given system of polynomial differential equations (or classes of equations with coefficients as symbolic parameters) with as many good properties (e.g. weak reversibility, reversibility, small deficiency, small number of linkage classes or reactions, mass conservation, etc.) as possible.

It has been shown that, under mass action kinetics, a necessary and sufficient condition for realizability is a sign structure in the system of polynomial differential equations [24] (see Lemma 1 below). Once we know that a system can arise from a mass action system, the question still lies: can the system arise from a mass action system with good properties? This is what this present work looks at. There exist similar works outside the realm of mass action kinetics. For example, [4] finds mass action systems that generate the same differential equation as a given S-system or generalized mass action system. The approach allows recent results from Chemical Reaction Network Theory to be applied in Biochemical Systems Theory.

There are several sources of motivation for this problem.

1. Having fitted a system of differential equations to data, one may wonder whether the obtained equations can be interpreted as the mass action type deterministic model of an appropriate reaction network.

---

*Department of Mathematics and Department of Biomolecular Chemistry, University of Wisconsin–Madison

†Department of Mathematics, San Jose State University, San Jose, CA 95192, USA

‡Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary

§Department of Mathematics and Computer Science, Freie Universität, Berlin, Germany

¶Department of Analysis, Budapest University of Technology and Economics and Laboratory for Chemical Kinetics, Eötvös Loránd University, Budapest, Hungary

‖Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706-1325, USA

2. There is an internal requirement within this branch of science: One should like to know as much as possible about the structure of differential equations that arise from modelling a chemical system.

3. Given a system of polynomial differential equations in any field of pure or applied mathematics, one may wish to have statements on stability or oscillations, similar to those offered by the Horn-Jackson Theorem [26], Zero Deficiency Theorem [21], Volpert's theorem [50], or the Global Attractor Conjecture, where several cases have been proven [3, 12, 23, 35][1]. Then it comes in handy to see that the system of differential equations of interest belongs to a well behaving class.

4. Lastly, results of formal reaction kinetics (to use an expression introduced in [5, 6]), e.g. on the existence of stationary points, may offer alternative methods for solving problems in algebraic geometry [17, 32]. For instance, one might be able to show the existence of positive roots of a polynomial if the system of polynomial equations is known to be the right hand side of the induced kinetic differential equation of a reversible or weakly reversible reaction network [7].

The structure of our paper is as follows. Section 2 introduces the essential concepts of reaction networks and mass action systems. Section 3 formulates the problem we are interested in, that of realizability of kinetic differential equations. Section 4 treats two special cases: finding realizations for *compartmental models* (defined later) and weakly reversible networks. Section 5 focuses on the general problem of realizability. We first review existing algorithms available. Then we outline several procedures that modify a reaction network while preserving the system of differential equations, including adding and removing vertices from the reaction graph. Section 6 explores the relation between weakly reversible and complex balanced realizations. Here we work with families of symbolic kinetic differential equations. In Section 7, we prove the geometric meaning of zero deficiency, and discuss when is a realization unique. We also pose several conjectures that may be of interest to the reader. An Appendix contains a very large number of enlightening examples. The present paper intends to be a review of the realizability problem, while containing some new results.

A few remarks on notation and the use of words are in order.

1. The set of vectors in $\mathbb{R}^M$ with strictly positive coordinates is denoted by $\mathbb{R}^M_{>0}$. Also

$$\mathbf{x}^{\mathbf{y}} := x_1^{y_1} x_2^{y_2} \cdots x_M^{y_M}$$

for any $x_i > 0$ and $y_i \geq 0$. If $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N)$, then

$$\mathbf{x}^{\mathbf{Y}} = (\mathbf{x}^{\mathbf{y}_1}, \mathbf{x}^{\mathbf{y}_2}, \ldots, \mathbf{x}^{\mathbf{y}_N})^T.$$

2. We use differential equation to mean a *system of ordinary differential equations.* In the same vein, we use *polynomial* to usually mean a system of polynomial equations, i.e. a vector-valued function on $\mathbb{R}^M_{>0}$. By *monomial* we mean a scalar-valued function on $\mathbb{R}^M_{>0}$.

3. We say a reaction network *induces* a differential equation, if we formulate the mass action type differential equation with given reaction rate coefficients, and a differential equation *is realized* by a reaction network, if the network induces the given differential equation.

2 Chemical reaction networks and mass action systems

Here we recapitulate some concepts of the mathematical description of reaction networks; for more details, see e.g. [14, 21, 49].

---

[1] A proof in full generality has been proposed in [9].

A **reaction network** consists of a set of R reaction steps among the chemical species X(1), X(2), ..., X(M):

$$\sum_{m=1}^{M} \alpha(m,r)X(m) \rightarrow \sum_{m=1}^{M} \beta(m,r)X(m) \quad (r = 1, 2, \ldots, R) \tag{1}$$
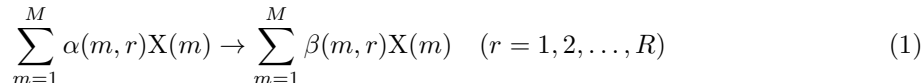
with nonnegative integer **stoichiometric coefficients** $\alpha(m,r)$, $\beta(m,r)$. Its induced kinetic differential equation assuming mass action type kinetics (and disregarding the change of temperature, pressure and reaction volume) is

$$\frac{dx_m}{dt} = \sum_{r=1}^{R} (\beta(m,r) - \alpha(m,r))k_r \prod_{p=1}^{M} x_p^{\alpha(p,r)} \quad (m = 1, 2, \ldots, M). \tag{2}$$

The positive number k_r is the **rate coefficient**[2] for the r-th reaction.

On the two sides of arrows in (1), which represent a reaction step, one has formal linear combinations of the species, called **complexes**, the left one being the **reactant complex**, and the right one being the **product complex**. Associated to each complex is the vector of the coefficients of these linear combinations. By an abuse of notation, we will refer to this vector as a complex. The difference between the product complex and the reactant complex is the **reaction vector**. These vectors form the columns of the **stoichiometric matrix**. The **stoichiometric space** $\mathcal{S}$ is the linear span of the reaction vectors, where $S = \dim \mathcal{S}$. (This number can also be interpreted as the number of independent reaction steps.) The **stoichiometric compatibility class** of $\mathbf{x}_0 \in \mathbb{R}_{>0}^M$ is the forward-invariant set $(\mathbf{x}_0 + \mathcal{S})_> := (\mathbf{x}_0 + \mathcal{S}) \cap \mathbb{R}_{>0}^M$. The number of complexes is denoted by N.

With sets of reactions and complexes as described above, each reaction network is naturally associated to a directed graph.

Definition 1. The **reaction graph** (or **Feinberg-Horn-Jackson graph**) of a given reaction network is a finite directed graph with no self-loops, where the vertices are the complexes of the reaction network, and the edges are precisely the reactions.

The connected components of a reaction graph are also called **linkage classes**, and L denotes the number of linkage classes in a reaction network. The **deficiency** of the reaction network is $\delta := N - L - S$, where N is the number of complexes (i.e., number of nodes in the reaction graph), and $S = \dim \mathcal{S}$. The strong components of the reaction graph are also called **strong linkage classes**. An **ergodic component** (or **strong terminal class**) is a strong component with no reaction step such that the reactant complex is in the component while the product complex is outside of it, i.e., an ergodic component is a strongly connected component of the reaction graph. The number of ergodic components will be denoted by T.

Definition 2. A reaction network is **reversible** if its reaction graph as a relation is symmetric, i.e., $i \rightarrow j$ is a reaction if and only if $j \rightarrow i$ is a reaction. It is **weakly reversible**, if the transitive closure of its reaction graph is symmetric; or if all the reaction arrows are edges of a directed cycle in the reaction graph; or all of its strong components are ergodic.[3]

Definition 3. A reversible reaction network (with rate coefficients k_{r+}, k_{r-} for the r-th pair of reversible reactions)

$$\sum_{m=1}^{M} \alpha(m,r)X(m) \rightleftharpoons \sum_{m=1}^{M} \beta(m,r)X(m) \quad (r = 1, 2, \ldots, P := R/2) \tag{3}$$

---
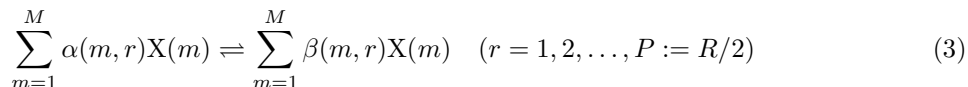
[2]Instead of *rate constants* we use the term *rate coefficients*, because they do depend on everything (pressure, temperature, volume), except concentrations, although we know that there is a heated argument about this question amongst reaction kineticists, see [31, p. 115].

[3]Weak reversibility implies $T = L$, but the converse is false [22].

is **detailed balanced** at the positive stationary concentration $\mathbf{x}^*$ if for each pair of reversible reactions, the forward step proceeds at the same rate as the backward step; in other words, $\mathbf{x}^*$ is detailed balanced if

$$k_{r+}(\mathbf{x}^*)^{\boldsymbol{\alpha}_r} = k_{r-}(\mathbf{x}^*)^{\boldsymbol{\beta}_r}, \quad (r = 1, 2, \ldots, P).$$

The induced kinetic differential equation can be written in many different forms, see e.g. [49, Sec. 6.3]. We introduce the form that will be used frequently below.

For a reaction graph with N complexes, let their associated vectors be columns of the matrix $\mathbf{Y} := (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N)$. Implicitly we are imposing an ordering on the set of complexes. For a reaction $i \to j$ (from the i-th complex to the jth complex), assume that the rate coefficient is $k_{ji} > 0$. Define the matrix $\mathbf{K}$ to be

$$[\mathbf{K}]_{ij} = \begin{cases} k_{ij} & \text{if } j \to i \text{ is a reaction in the graph} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

so that $\mathbf{K}$ is a matrix with nonnegative entries, and 0 along the diagonal. Then the column-conserving matrix $\mathbf{A}_k := \mathbf{K} - diag(\mathbf{K}^\top \mathbf{1})$ is the **(weighted) Laplacian** of the reaction graph. A matrix with nonnegative off-diagonal entries, and is column-conserving is also called **Kirchoff**.

With these matrices defined, the induced kinetic differential equation (2) of the reaction network can be rewritten in the following form:

$$\frac{d\mathbf{x}}{dt} = \mathbf{Y} \cdot \mathbf{A}_k \cdot \mathbf{x}^{\mathbf{Y}}, \tag{5}$$

where $\mathbf{x}^{\mathbf{Y}}$ is the vector of monomials $(\mathbf{x}^{\mathbf{y}_1}, \mathbf{x}^{\mathbf{y}_2}, \ldots, \mathbf{x}^{\mathbf{y}_N})^T$.

Note that the differential equation (5) is completely determined by the complexes (columns of $\mathbf{Y}$) and the connectivity and rate coefficients of the reaction graph (given by $\mathbf{A}_k$).

Definition 4. The reaction system (5) is **complex balanced** at the positive stationary concentration $\mathbf{x}^*$, if all the complexes are destroyed and formed with the same rate at this concentration:

$$\sum_{q \neq n} k_{qn}(\mathbf{x}^*)^{\mathbf{y}_n} = \sum_{q \neq n} k_{nq}(\mathbf{x}^*)^{\mathbf{y}_q}$$

for all $n = 1, 2, \ldots, N$. Equivalently, $\mathbf{x}^*$ is complex balanced if $\mathbf{A}_k(\mathbf{x}^*)^{\mathbf{Y}} = \mathbf{0}$.

If one remains within the framework of mass action kinetics – as we do here – then the Deficiency Zero Theorem provides a necessary and sufficient structural condition for complex balancing [19, 25, 26]:

Theorem 1. A mass action system is complex balanced for all choices of the rate coefficients if and only if it is weakly reversible and its deficiency is zero.

One might notice that in the Deficiency Zero Theorem, it is the mass action system that is described as complex balanced, not one of its stationary concentrations. This is because complex balancing implies a host of good structural and dynamical properties, including:

1. If one of the stationary concentrations is complex balanced, then all stationary concentrations of the system is complex balanced.

2. There is exactly one stationary concentration within every stoichiometric compatibility class.

3. Let $\mathbf{x}^*$ be any complex balanced stationary concentration. The function

$$L(\mathbf{x}) := \sum_{m=1}^{M} (\ln x_m - \ln x_m^* - 1)$$

defined on $\mathbb{R}_{>0}^n$ is a Lyapunov function, with global minimum at $\mathbf{x} = \mathbf{x}^*$.

4. Every positive stationary concentration is locally asymptotically stable within its stoichiometric compatibility class.

5. The set of complex balanced stationary concentrations $Z_{\mathbf{k}}$ satisfies the equation $\ln Z_{\mathbf{k}} = \ln \mathbf{x}^* + \mathcal{S}^\perp$.

3   Formulation of problems

In defining the system of differential equations for a mass action system, note that given a set of reactions and rate coefficients (or equivalently, given a reaction graph and positive weights for each edge), it is straightforward to write down the differential equation. Unfortunately, given a differential equation (satisfying a mild sign condition introduced below), there is in general no unique reaction graph that is associated to it [13].

A natural question arises: given a differential equation, what are the possible reaction graphs that could generate it if we assume mass action kinetics. All the information about a reaction graph is contained inside $(\mathbf{Y}, \mathbf{A}_k)$.

Definition 5. A **realization** of a differential equation (5) is the pair $(\mathbf{Y}, \mathbf{A}_k)$, where $\mathbf{Y} \in \mathbb{Z}_{\geq 0}^{M \times N}$ and $\mathbf{A}_k$ is the weighted Laplacian of a reaction graph. Equivalently, a realization is a set of complexes and a Feinberg-Horn-Jackson reaction graph. Two different realizations $(\mathbf{Y}, \mathbf{A}_k)$ and $(\mathbf{Y}', \mathbf{A}'_k)$ are said to be **dynamically equivalent** if they give rise to the same differential equation under mass action kinetics.

Dynamical equivalence is written as linear relations in [11]. In other words, the question we are asking in this work is: what are the possible dynamically equivalent realizations $(\mathbf{Y}, \mathbf{A}_k)$ of a given system of differential equations?

Before searching for dynamically equivalent realizations, we must first ask whether or not a given differential equation can arise from mass action kinetics. In other words, is it even possible to find one realization? We provide a necessary and sufficient condition for realizability in Lemma 1 below.

Assume that we are given a system of differential equations of the form

$$\dot{\mathbf{x}} = \mathbf{Z} \cdot \mathbf{x}^{\mathbf{Y}} = \sum_{n=1}^N \mathbf{z}_n \mathbf{x}^{\mathbf{y}_n} = \begin{pmatrix} \sum_{n=1}^N z_n^1 \mathbf{x}^{\mathbf{y}_n} \\ \vdots \\ \sum_{n=1}^N z_n^m \mathbf{x}^{\mathbf{y}_n} \\ \vdots \\ \sum_{n=1}^N z_n^M \mathbf{x}^{\mathbf{y}_n} \end{pmatrix}, \tag{6}$$

where $\mathbf{Z} := (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N)$, and $\mathbf{Y} := (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N)$, and $\mathbf{x}^{\mathbf{y}} := x_1^{\mathbf{y}_1} x_2^{\mathbf{y}_2} \cdots x_M^{\mathbf{y}_M}$ for any $x_i > 0$.

Definition 6. We say that a system of differential equations (6) is **kinetic** if

$$z_n^m < 0 \quad \text{implies} \quad (\mathbf{y}_n)^m = y_n^m > 0. \tag{7}$$

A system being kinetic reflects the physical assumption that in order to lose a chemical species via a reaction, that species must participate as a reactant. For example, the system

$$\dot{x} = y - x$$
$$\dot{y} = x - xz - y$$
$$\dot{z} = xy - z$$

is not kinetic because the term $-xz$ in $\dot{y}$ does not involve the variable y.

It turns out that a differential equation can arise from mass action kinetics if and only if it is **kinetic** in the sense of Definition 6. This has been shown in [24] using a constructive method.

Lemma 1 (Hungarian Lemma)**.** A polynomial system of differential equations of the form (6) can be realized by a mass action system if and only if the system of differential equations is kinetic.

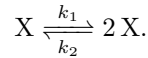One realization (the *canonical realization*) is the collection of reactions

$$\mathbf{y}_n \xrightarrow{|z_n^m|} \mathbf{y}_n + \text{sign}(z_n^m)\mathbf{e}_m \quad (m = 1, 2, \ldots, M; n = 1, 2, \ldots, N), \tag{8}$$
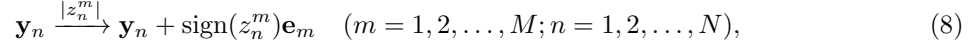
where $\mathbf{e}_m$ is the m-th vector of the standard basis of $\mathbb{R}^M$. Here the product complex vector has nonnegative integer coordinates because of the assumption in (7).

Example 1. Given the differential equation

$$\dot{x} = k_1 x - k_2 x^2 \tag{9}$$

(with k_1, $k_2 > 0$), it is possible to find an inducing reaction network with only two complexes defined by the exponents of the monomials present in the right hand side:

$$\text{X} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 2\,\text{X}.$$

The right hand side contains two monomials, and there exists a realization with two complexes, and is also reversible, and has deficiency zero. This also happens to be the canonical realization.

However, in general, the canonical realization is quite complicated, contains too many complexes and reactions (see Supplementary Example 13). Its main advantage is that it can be automatically constructed and used as a starting point for finding different realizations.

Now we present the set of problems we are interested in (also posed earlier in [18, Section 4.7]).

List of Problems.

1. Given a kinetic differential equation $\dot{\mathbf{x}} = \mathbf{Z} \cdot \mathbf{x}^\mathbf{Y}$, find $\mathbf{A}_k$ such that $\mathbf{Y} \cdot \mathbf{A}_k = \mathbf{Z}$.

2. Find reversible or weakly reversible; complex balanced or detailed balanced; or mass conserving realizations. Decide if such a realization exists.

3. Find realizations with minimal deficiency, with minimal number of reactions, or with minimal number of complexes.

4. Without further restrictions, the question about uniqueness of $\mathbf{A}_k$ arises at all levels.

5. In all of the above problems, we may add more columns to $\mathbf{Y}$ [41]. For example, if the first problem has no solution, we may consider the following: find $\overline{\mathbf{Y}}$ whose first N columns form precisely $\mathbf{Y}$ and find $\mathbf{A}_k$ such that $\overline{\mathbf{Y}} \cdot \mathbf{A}_k \cdot \mathbf{x}^{\overline{\mathbf{Y}}} = \mathbf{Z} \cdot \mathbf{x}^\mathbf{Y}$. In this case we can also ask what is the minimum number of columns that must be added.

We will answer partially the questions above, especially pertaining to reversible, weakly reversible, complex balanced and detailed balanced realizations. We also consider uniqueness in Section 7 and mass conserving realizations in the Supplementary.

In considering the uniqueness question, the concepts of sparse and dense realizations have been useful [42]. The aim is to find an $M \times N$ matrix of nonnegative integer components $\mathbf{Y}$ and a Kirchoff matrix $\mathbf{A}_k$ giving the product $\mathbf{Z} = \mathbf{Y} \cdot \mathbf{A}_k$. The pair $(\mathbf{Y}, \mathbf{A}_k)$ is a realization of the matrix $\mathbf{Z}$. Since $(\mathbf{Y}, \mathbf{A}_k)$ uniquely identifies the reaction graph, when we refer to a subgraph of the realization, we mean a subgraph of the reaction graph corresponding to $(\mathbf{Y}, \mathbf{A}_k)$.

Definition 7. Suppose the set of complexes is fixed, i.e., the matrix $\mathbf{Y}$ is given. If the number of zeros in the above setting in $\mathbf{A}_k$ is maximal, then a ***sparse realization*** has been obtained. If the number of zeros in $\mathbf{A}_k$ is minimal then a ***dense realization*** has been obtained.

Suppose $\mathbf{Z}$ is given and fix the matrix $\mathbf{Y}$, i.e., fix the set of complexes (both the reactant and product complexes) from which we will search for realizations. The following properties of dense and sparse realizations of a given chemical reaction network were proved in [45]:

Theorem 2.

1. The reaction graph of the dense realizations is unique.

2. The reaction graph of any dynamically equivalent realization is a subgraph of the dense realization.

3. The reaction graph of a system of differential equations is unique if and only if the reaction graph of the dense and sparse realizations are identical.

The above statements hold when the word "realization" is replaced by *weakly reversible realization* [1]. The above results were extended to the case of *constrained* realizations, where the rate coefficients fulfil arbitrary polytopic constraints [2]. An important special case of this is when a subset of possible reactions is excluded from the network.

We have seen that the canonical representation may be the sparse realization. Later, we shall give a complete solution to a generalization of part 1 of Problem 3 for the case when the numerical values of the components of the matrices $\mathbf{Z}$ and $\mathbf{Y}$ are known. We also review results for the solution of the other related problems in Section 5.1.

To date, the problem of finding a realization has used tools and techniques from different mathematical disciplines, ranging from graph theory, combinatorics, linear algebra, to optimization.

# 4 Realizations of compartmental models and reversible systems

In this section, we consider two classes of examples: compartmental models and weakly reversible networks. Because of the simpler mathematical structure of compartmental models, the realization problem can be solved even if the rate coefficients are unspecified parameters, i.e., symbolic. Then we will look at an *ad hoc* method to find reversible realizations.

4.1 Compartmental Models

A ***compartmental system*** is built using a subset of reactions of the form
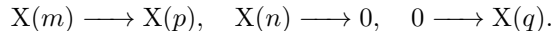
$$\mathrm{X}(m) \longrightarrow \mathrm{X}(p), \quad \mathrm{X}(n) \longrightarrow 0, \quad 0 \longrightarrow \mathrm{X}(q).$$

It is ***closed*** if it only contains reactions of the first type, it is ***half-open*** if reactions of the second type are also allowed, and it is ***open*** if reactions of the third type may occur. A ***generalized compartmental system*** only consists of the reactions

$$y_m \mathrm{X}(m) \longrightarrow y_p \mathrm{X}(p), \quad y_n \mathrm{X}(n) \longrightarrow 0, \quad 0 \longrightarrow y_q \mathrm{X}(q)$$
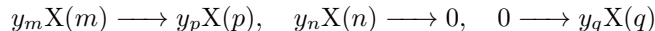
with fixed positive integer numbers $y_1, y_2, \ldots, y_N$. (It is understood that all the complexes contain a single species, and every species appear in a single complex.) It is ***closed***, if it only contains reactions of the first type, it is ***half-open*** if reactions of the second type are also allowed, and it is ***open***, if reactions of the third type also may occur.

First we give a complete, although ineffective solution of the first task of Problem 3 in a slightly modified form. Modification means that the empty complex will be dealt with separately, because in some calculations this implies simplification. First, we need two definitions.

Definition 8. [47, pp. 142–143] A matrix $\mathbf{B} \in \mathbb{R}^{N \times N}$ is a ***compartmental matrix***, if its off-diagonal elements are nonnegative, while its diagonal elements are smaller than or equal to the negative of the corresponding column sums: $b_{nn} \leq -\sum_{p=1}^{N} b_{pn}$. It is a ***Kirchhoff matrix*** if it is a compartmental matrix and its diagonal elements are equal to the negative of the corresponding column sums: $b_{nn} = -\sum_{p \neq n} b_{pn}$.

For example, the weighted Laplacian matrix of a reaction graph is Kirchoff.

Let $M, N \in \mathbb{N}$; and suppose the vectors $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N \in \mathbb{Z}_{\geq 0}^M$ are arbitrary nonzero vectors. Let us write the reaction network we are interested in, in the form of

$$\sum_{m=1}^{M} y_q^m \mathrm{X}(m) \to \sum_{m=1}^{M} y_n^m \mathrm{X}(m) \quad (n, q = 1, 2, \ldots, N; n \neq q), \tag{10}$$

$$\sum_{m=1}^{M} y_q^m \mathrm{X}(m) \to 0 \quad (q = 1, 2, \ldots, N), \tag{11}$$

$$0 \to \sum_{m=1}^{M} y_n^m \mathrm{X}(m) \quad (n = 1, 2, \ldots, N). \tag{12}$$

Let us introduce the notation (slightly different from the usual)

$$\mathbf{K} := \left(\ k_{nq} \ \right)_{n \neq q=1}^{N}, \quad \mathbf{u} := \begin{pmatrix} k_{01} \\ k_{02} \\ \vdots \\ k_{0N} \end{pmatrix}, \quad \mathbf{v} := \begin{pmatrix} k_{10} \\ k_{20} \\ \vdots \\ k_{N0} \end{pmatrix},$$

where $\mathbf{A}_k := \mathbf{K} - diag(\mathbf{K}^T \mathbf{1})$ is the weighted Laplacian of the reaction network. Note that $k_{nq} \geq 0$. Then, the induced kinetic differential equation of the reaction network (10)–(12) is

$$\dot{\mathbf{c}} = \mathbf{Y}(\mathbf{K} - diag(\mathbf{K}^\top \mathbf{1} - \mathbf{u}))\mathbf{c}^{\mathbf{Y}} + \mathbf{Yv} \quad \text{with } (\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N)). \tag{13}$$

Upon denoting $\mathbf{Z} := \mathbf{Y}(\mathbf{K} - diag(\mathbf{K}^\top \mathbf{1} - \mathbf{u})), \mathbf{b} := \mathbf{Yv}$ one can say that (13) is of the form

$$\dot{\mathbf{c}} = \mathbf{Zc}^{\mathbf{Y}} + \mathbf{b} \tag{14}$$

so that there exists a compartmental matrix $\mathbf{B} \in \mathbb{R}^{N \times N}$ and a nonnegative vector $\mathbf{v} \in \mathbb{R}_{\geq 0}^M$ so that

$$\mathbf{Z} = \mathbf{YB}, \quad \mathbf{b} = \mathbf{Yv} \tag{15}$$

hold. If $\mathbf{u} = \mathbf{0}$, then $\mathbf{B}$ is a Kirchhoff matrix. (One may call (11) *outflow*, (12) *inflow*, and they are missing if $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$, respectively.)

A partial answer to the realization problem follows.

Theorem 3. Suppose that the differential equation

$$\dot{\mathbf{x}} = \mathbf{Zx}^{\mathbf{Y}} + \mathbf{b} \tag{16}$$

where $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N), \mathbf{Z} \in \mathbb{R}^{M \times N}$ (with $M, N \in \mathbb{N}$) has the property that there exist a compartmental matrix $\mathbf{B}$ and a nonnegative vector $\mathbf{v} \in \mathbb{R}_{\geq 0}^M$ so that (15) holds. Then the equation has a realization of the form (10)–(12).

Proof. First we prove that the conditions of the theorem imply the differential equation is kinetic. Suppose $Z_n^m < 0$ for some indices m, n. Then,

$$Z_n^m = \sum_{q=1}^{N} y_q^m B_n^q = \sum_{q \neq n} y_q^m B_n^q + y_n^m B_n^n,$$

and the last sum being nonnegative, $y_n^m B_n^n$ should be negative, thus $y_n^m > 0$.

The components of $\mathbf{Y}$ and $\mathbf{v}$ are nonnegative, thus the components of $\mathbf{b} = \mathbf{Yv}$ are nonnegative, as well.

Next, we construct the inducing realization in the following way. Let the reaction rate coefficients of the reaction network (10)–(12) be defined as follows:

$$
\begin{aligned}
k_{qn} &:= B_n^q \quad (n, q = 1, 2, \ldots, N; n \neq q), \\
k_{n0} &:= v_n \quad (n = 1, 2, \ldots, N), \\
k_{0n} &:= -B_n^n - \sum_{q=1}^{N} B_n^q \quad (n = 1, 2, \ldots, N).
\end{aligned}
$$

Then simple verification proves the statement. $\qquad\square$

It may happen that a given kinetic differential equation has zero, one or an infinite number of realizations with complexes, or columns of $\mathbf{Y}$, determined by the exponents of its monomials, see Supplementary Examples 11, 12 and 15 and Example 2, respectively.

For some nonlinear differential equations one can provide a zero deficiency realization. Additional conditions may allow to have a realization which is also reversible or weakly reversible. One of the early approaches to this problem was [48, Theorem 9], see also [49, Theorem 11.12] providing a necessary and sufficient condition that a generalized compartmental system induces a given differential equation. If the coefficients of the kinetic polynomial ODEs are known, then a deficiency zero realization can be determined via mixed integer linear programming, if it exists [33]. Moreover, weakly reversible deficiency zero realizations can be determined using simple linear programming [43].

Theorem 4. Let $M \in \mathbb{N}; \mathbf{Z} \in \mathbb{R}^{M \times M}, \mathbf{Y} \in \mathbb{N}^{M \times M}, \mathbf{b} \in \mathbb{R}_{\geq 0}^{M}$. The differential equation

$$
\dot{\mathbf{x}} = \mathbf{Z}\mathbf{x}^{\mathbf{Y}} + \mathbf{b} \tag{17}
$$

is the induced kinetic differential equation of a

1. closed,

2. half open,

3. open

generalized compartmental system if and only if the following relations hold: all the columns of $\mathbf{Y}$ are positive multiples of different elements of the standard basis:

$$
\mathbf{Y} = \begin{bmatrix} y_1\mathbf{e}_1 & y_2\mathbf{e}_2 & \ldots & y_M\mathbf{e}_M \end{bmatrix}; y_1, y_2, \ldots, y_M \in \mathbb{N};
$$

and

1. $b_m = 0; -z_m^m, z_p^m, \beta_m := \sum_{p=1}^{M} \frac{z_m^p}{y_p} \geq 0; z_m^m = -\beta_m y_m;$

2. $b_m = 0; -z_m^m, z_p^m, \beta_m \geq 0; z_m^m \leq -\beta_m y_m, \exists m : z_m^m < -\beta_m y_m;$

3. $-z_m^m, z_p^m, \beta_m \geq 0; z_m^m \leq -\beta_m y_m, \exists m : b_m > 0;$

respectively, where throughout $m, p \in \{1, 2, \ldots, M\}, m \neq p$.

Proof. Necessity is obvious. Instead of direct calculations sufficiency can be learned from the application of Theorem 3, as in this case $\mathbf{Y}$ – being diagonal with positive components – is invertible, and its inverse can easily be calculated. $\qquad\square$

Remark 1.

1. A necessary and sufficient condition of reversibility in the closed case is sign symmetry of the coefficient matrix. Similar conditions can be formulated for the other two cases, as well.

2. Once we have a zero deficiency reversible realization, it is easy to check detailed balancing: only the circuit conditions [20] are to be checked.

4.2 Finding reversible realizations

Let us study reversibility, another relevant property, in more detail, applying different methods— without taking care of the value of deficiency.

An equivalent and useful characterization of the induced kinetic differential equation of reversible reaction networks seems to be missing. What we only have is a heuristic algorithm without the proof that it leads to a reversible realization if there exists one, and the realization is independent from the order of the steps.

Example 2.

1. Consider the polynomial equation

$$\dot{x} = b + s - gx + sz - gxz \tag{18}$$
$$\dot{y} = b + s + sx - gy - gxy \tag{19}$$
$$\dot{z} = b + s + sy - gz - gyz \tag{20}$$

with $b, g, s > 0$ (a transform of the first repressilator model in [17]). The canonical representation of this equation is shown in Figure 1. It is a reversible reaction network with exactly those seven complexes which are given as exponents of the monomials. It consists of a single linkage class and has a three-dimensional stoichiometric space, and it is also a sparse realization. Reversibility allows us to deduce the existence of a positive stationary point by the theorems of [7, 34]. In the third part below a more systematic approach to this example will show an infinite number of realizations (not necessarily reversible), actually we shall receive the dense realization.
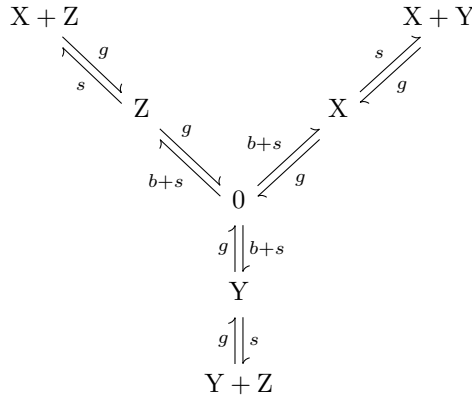

Figure 1: A reversible single linkage class reaction network of deficiency 3 to induce Eq. (18)-(20). This realization is sparse as opposed to the dense realization shown in Figure 2.

2. To find a reversible realization one can also use a heuristic: one tries to collect pairs of reactant and product complex vectors. Suppose our goal is to find a reversible reaction network inducing (18)–(20). Let us rewrite it in the following form

$$\begin{pmatrix} b + s - gx + sz - gxz \\ b + s + sx - gy - gxy \\ b + s + sy - gz - gyz \end{pmatrix} = gxz \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + \dots. \tag{21}$$

Now we consider the exponent of the monomial as a reactant vector $\boldsymbol{\alpha}$ and the vector with which this monomial has been multiplied, as the reaction vector $\boldsymbol{\beta} - \boldsymbol{\alpha}$, then one should have

$$\boldsymbol{\beta} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$ thus we need a term where the exponent of the monomial

is $\boldsymbol{\beta}$, and it is multiplied by a positive constant and $\boldsymbol{\alpha} - \boldsymbol{\beta}$. Thus we get:

$$\begin{pmatrix} b + s - gx + sz - gxz \\ b + s + sx - gy - gxy \\ b + s + sy - gz - gyz \end{pmatrix} = gxz \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + sz \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \dots. \tag{22}$$

Repeating steps of this kind we arrive at the decomposition

$$gxz \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + sz \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + gyz \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + sy \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} +$$

$$gxy \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + sx \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + gx \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + (b+s) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} +$$

$$gy \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + (b+s) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + gz \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + (b+s) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

which again leads to the network in Figure 1.

Conjecture 1. If there exists a reversible realization then the above process (independently from the order of actions) ends and necessarily provides one of them which may or may not depend on the order of steps.

3. Finally, one can apply Theorem 3 to get all the realizations (with the exponents as complexes only) and choose from them the reversible one(s). Suppose again that the Eqs. (18)–(20) are given, and our goal is to find a reaction network inducing it using Theorem 3. In this case we have

$$\mathbf{Z} = \begin{pmatrix} -g & 0 & s & -g & 0 & 0 \\ s & -g & 0 & 0 & -g & 0 \\ 0 & s & -g & 0 & 0 & -g \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b+s \\ b+s \\ b+s \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix} \tag{23}$$

and we have to find a compartmental solution in $\mathbf{B}$ of the equation $\mathbf{Z} = \mathbf{Y}\mathbf{B}$, or

$$\begin{array}{rclcrclcrcl}
-g & = & -b_{21} - b_{31} - b_{61} - u_1 & \quad s & = & b_{21} + b_{51} + b_{61} & \quad 0 & = & b_{31} + b_{41} + b_{61} \\
0 & = & b_{12} + b_{42} + b_{52} & \quad -g & = & -b_{12} - b_{32} - b_{42} - u_2 & \quad s & = & b_{32} + b_{42} + b_{62} \\
s & = & b_{13} + b_{43} + b_{53} & \quad 0 & = & b_{23} + b_{53} + b_{63} & \quad -g & = & -b_{13} - b_{23} - b_{53} - u_3 \\
-g & = & -b_{24} - b_{34} - b_{64} - u_4 & \quad 0 & = & b_{24} + b_{54} + b_{64} & \quad 0 & = & -b_{14} - b_{24} - b_{54} - u_4 \\
0 & = & -b_{25} - b_{35} - b_{65} - u_5 & \quad -g & = & -b_{15} - b_{35} - b_{45} - u_5 & \quad 0 & = & b_{35} + b_{45} + b_{65} \\
0 & = & b_{16} + b_{46} + b_{56} & \quad 0 & = & -b_{16} - b_{36} - b_{46} - u_6 & \quad -g & = & -b_{16} - b_{26} - b_{56} - u_6.
\end{array}$$

11

which for the coordinates has the following consequences.

$$b_{31} = 0 \quad b_{41} = 0 \quad b_{61} = 0 \quad b_{12} = 0 \quad b_{42} = 0 \quad b_{52} = 0 \quad b_{23} = 0 \quad b_{53} = 0 \quad b_{63} = 0$$
$$b_{24} = 0 \quad b_{54} = 0 \quad b_{64} = 0 \quad b_{35} = 0 \quad b_{45} = 0 \quad b_{65} = 0 \quad b_{16} = 0 \quad b_{46} = 0 \quad b_{56} = 0$$

and then the equations reduce to

$$
\begin{array}{rclcrclcrclcrcl}
-g &=& -b_{21} - u_1 & \quad & s &=& b_{21} + b_{51} & \quad & -g &=& -b_{32} - u_2 & \quad & s &=& b_{32} + b_{62} \\
s &=& b_{13} + b_{43} & \quad & -g &=& -b_{13} - u_3 & \quad & -g &=& -b_{34} - u_4 & \quad & 0 &=& -b_{14} - u_4 \\
0 &=& -b_{25} - u_5 & \quad & -g &=& -b_{15} - u_5 & \quad & 0 &=& -b_{36} - u_6 & \quad & -g &=& -b_{26} - u_6
\end{array}
$$

having the solution

$$
\begin{array}{rclcrclcrclcrcl}
b_{21} &=& g - u_1 & \quad b_{51} &=& s - g + u_1 & \quad b_{32} &=& g - u_2 & \quad b_{62} &=& s - g + u_2 \\
b_{13} &=& g - u_3 & \quad b_{43} &=& s - g + u_3 & \quad b_{34} &=& g - u_4 & \quad b_{14} &=& -u_4 \\
b_{25} &=& -u_5 & \quad b_{15} &=& g - u_5 & \quad b_{36} &=& -u_6 & \quad b_{26} &=& g - u_6
\end{array}
$$

implying that

$$u_4 = u_5 = u_6 = 0.$$

Now the kinetic matrix of the dense realization (without inflow) is

$$
\begin{pmatrix}
-s - u_1 & 0 & g - u_3 & 0 & g & 0 \\
g - u_1 & -s - u_2 & 0 & 0 & 0 & g \\
0 & g - u_2 & -s - u_3 & g & 0 & 0 \\
0 & 0 & -g + s + u_3 & -g & 0 & 0 \\
-g + s + u_1 & 0 & 0 & 0 & -g & 0 \\
0 & -g + s + u_2 & 0 & 0 & 0 & -g
\end{pmatrix}
$$

under the conditions that

$$s - g \leq u_1 \leq g \quad s - g \leq u_2 \leq g \quad s - g \leq u_3 \leq g$$

hold. To have no indices let us introduce $\alpha := u_1, \beta := u_2, \gamma := u_3$. Now consider two cases:

(a) If $g \geq s$, take α, β and γ such that $0 \leq \alpha, \beta, \gamma \leq s$.

(b) If $s \geq g$, take α, β and γ such that $s - g \leq \alpha, \beta, \gamma \leq s$.

As to the inflow: the components of the solution $\mathbf{v}$ to $\mathbf{Yv} = \mathbf{b}$ should fulfil

$$v_1 + v_4 + v_5 = b + s \quad v_2 + v_5 + v_6 = b + s \quad v_3 + v_4 + v_6 = b + s,$$

which means that any $\mathbf{v}$ defines a solution for which

$$0 \leq v_4, 0 \leq v_5, 0 \leq v_6, v_4 + v_5 \leq b + s, v_5 + v_6 \leq b + s, v_4 + v_6 \leq b + s$$

hold; and then $v_1 := b + s - v_4 - v_5, v_2 := b + s - v_5 - v_6, v_3 := b + s - v_4 - v_6$.

Summing up, Figure 1 shows a sparse realization, while Figure 2 shows the dense realization. We also obtain the condition of reversibility: $\alpha = \beta = \gamma = s$ and $v_4 = v_5 = v_6 = 0$, and that it is exactly the sparse realization which is reversible.

In this example we were lucky to receive a canonical realization with so many nice properties. The reader can easily verify that this is the case with the second three dimensional repressilator model and with the six dimensional one of [17]: the canonic realization is again reversible thus assures the existence of a positive stationary point.

Let us emphasize that via a known result in reaction kinetics we were able to assert the existence of a positive solution to a second degree polynomial (in three variables), or to put it another way, we could *apply formal reaction kinetics in mathematics*.

As to uniqueness of the stationary point: it does not follow from the mentioned theorems. One may try several methods, see e.g. [30]. However, uniqueness of the reversible realization itself does follow (in this special case) from the method based on Theorem 3.
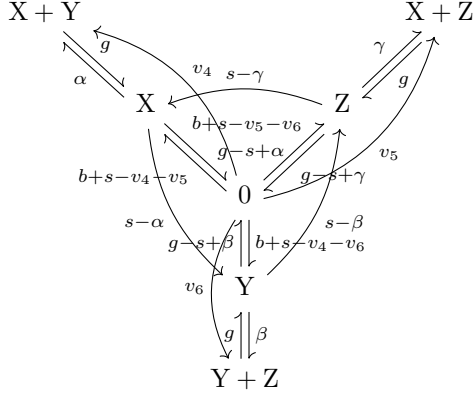
Figure 2: A weakly reversible single linkage class realization of deficiency 3 to induce Eq. (18)-(20). This realization is dense as opposed to the sparse realization shown in Figure 1.

5 On computing realizations of kinetic equations

In the previous section, we looked at two classes of systems with certain network properties, namely that of (generalized) compartmental models (i.e., complexes consisting of a single species) and that of reversible networks. In this section, we will provide an algorithmic approach to the realizability problem when the coefficients in the ODE are given. To date, there are algorithms based on linear programming (LP) and mixed integer linear programming (MILP) techniques. These are reviewed in Section 5.1.

In the previous section, we only use the complexes that show up as exponents of monomials in the kinetic differential equation. This may not necessarily be the case for more general examples [41]. However, once *a set of complexes has been decided on* (and must include the exponents of monomials), the algorithms presented below can be used to compute realizations with desirable structural properties.

In Section 5.2, we consider ways to modify an existing realization while preserving dynamical equivalence, including adding to/subtracting from the set of complexes. Finally, we consider cases when a realization with certain properties (e.g., complex balanced) is independent of the choice of the set of complexes.

5.1 Review of algorithms

Qualitative properties can be rewritten in terms of a *mixed integer linear programming* (MILP) problem [45]. The works [27, 46] determine weakly reversible realizations, while [43] provides reversible and detailed balanced realizations. The paper [40] finds detailed balanced subnetworks in reactions.

Recall the discussion about possibly enlarging the set of complexes (Problem 5) from $\mathbf{Y}$ to $\overline{\mathbf{Y}}$. We write the dynamical equivalence conditions for finding a realization using linear programming method with this predetermined complex set $\overline{\mathbf{Y}}$. Note that the columns of $\mathbf{Y}$ must also be columns of $\overline{\mathbf{Y}}$. We want to find a realization of the kinetic differential equation $\dot{\mathbf{x}} = \mathbf{Z} \cdot \mathbf{x}^{\mathbf{Y}}$ with the complex set $\overline{\mathbf{Y}}$. In other words, we want to find a Kirchoff matrix $\mathbf{A}_k$ such that

$$\overline{\mathbf{Y}} \cdot \mathbf{A}_k \cdot \mathbf{x}^{\overline{\mathbf{Y}}} = \mathbf{Z} \cdot \mathbf{x}^{\mathbf{Y}}. \tag{24}$$

We may add columns of $\mathbf{0}$ to the matrix $\mathbf{Z}$ and obtain $\overline{\mathbf{Z}}$, where the number of columns added is the number of additional columns in $\overline{\mathbf{Y}}$ compared to $\mathbf{Y}$. Then Eq. (24) becomes

$$\overline{\mathbf{Y}} \cdot \mathbf{A}_k \cdot \mathbf{x}^{\overline{\mathbf{Y}}} = \overline{\mathbf{Z}} \cdot \mathbf{x}^{\overline{\mathbf{Y}}}. \tag{25}$$

Therefore, kinetic realizability using $\overline{\mathbf{Y}}$ is equivalent to the feasibility of the following convex constraint set

$$\overline{\mathbf{Y}} \cdot \mathbf{A}_k = \overline{\mathbf{Z}} \qquad \text{(dynamical equivalence)} \qquad (26)$$
$$[\mathbf{A}_k]_{ij} \geq 0 \text{ for } i \neq j \qquad \text{(nonnegativity of rate coefficients)} \qquad (27)$$
$$\mathbf{1}^\top \cdot \mathbf{A}_k = \mathbf{0} \qquad \text{(column conservation in } \mathbf{A}_k) \qquad (28)$$

Next, we review algorithms for finding realizations with certain structural properties, namely, reversible, weakly reversible, or complex balancing realizations. In Eq. (26)-(28) above, the matrices $\mathbf{Y}$ and $\mathbf{Z}$ do not show up explicitly. Hence, to simplify notation, in what follows below, we let $\mathbf{Y}$ denote the complexes to be considered, some of which may not appear explicitly in the differential equation, and $\mathbf{Z}$ defines a kinetic differential equation (with possibly columns of $\mathbf{0}$ so that $\mathbf{Z}$ is the same size as $\mathbf{Y}$).

To find a **reversible realization**, consider the following feasibility problem [28]:

$$\mathbf{Y} \cdot \mathbf{A}_k = \mathbf{Z} \qquad (29)$$
$$[\mathbf{A}_k]_{ij} \geq 0 \text{ for } i \neq j \qquad (30)$$
$$\mathbf{1}^T \cdot \mathbf{A}_k = \mathbf{0} \qquad (31)$$
$$[\mathbf{A}_k]_{ij} \geq \varepsilon \iff [\mathbf{A}_k]_{ji} \geq \varepsilon \text{ for } i \neq j. \qquad (32)$$

The statement (32) is the condition that the reaction step $i \to j$ is present if and only if reaction $j \to i$ is present. This feasibility problem can be solved in the framework of mixed integer linear programming (MILP).

To find a **weakly reversible realization**, consider the following feasibility problem [29]:

$$\mathbf{Y} \cdot \mathbf{A}_k = \mathbf{Z} \qquad (33)$$
$$[\mathbf{A}_k]_{ij} \geq 0 \text{ for } i \neq j \qquad (34)$$
$$\mathbf{1}^T \cdot \mathbf{A}_k = \mathbf{0} \qquad (35)$$
$$\varepsilon[\mathbf{A}_k]_{ij} \leq [\mathbf{A}'_k]_{ij}, \ \varepsilon[\mathbf{A}'_k]_{ij} \leq [\mathbf{A}_k]_{ij} \text{ for } i \neq j \qquad (36)$$
$$\mathbf{A}'_k \cdot \mathbf{1} = \mathbf{0}. \qquad (37)$$

The inequalities (36) ensure that the structure of $\mathbf{A}_k$ and $\mathbf{A}'_k$ are the same, i.e., a reaction is present in one if and only if it is present in the other. Since every weakly reversible network is complex balanced for some choice of parameter, the algorithm searches for a complex balanced $\mathbf{A}'_k$ (in Eq. (37)), which shares the network structure we want but there is no relations in terms of the rate coefficients we seek.

In this MILP problem, the decision variables are the off-diagonal entries of $\mathbf{A}_k$, the off-diagonals entries of another Kirchoff matrix $\mathbf{A}'_k$, while $\mathbf{Z}, \mathbf{Y}$ are given matrices, and $\varepsilon \ll 1$ is a fixed small constant. We will explore the relationship between finding weakly reversible and complex balanced realizations in Section 6. More efficient methods to compute weakly reversible realizations are available; for example, a polynomial time algorithm can be found in [38].

To find a **detailed balanced realization**, supposed we have the numerical value of one stationary concentration $\mathbf{x}^*$ and consider [43]:

$$\mathbf{Y} \cdot \mathbf{A}_k = \mathbf{Z} \qquad (38)$$
$$[\mathbf{A}_k]_{ij} \geq 0 \text{ for } i \neq j \qquad (39)$$
$$\mathbf{1}^T \cdot \mathbf{A}_k = \mathbf{0} \qquad (40)$$

$$[\mathbf{A}_k]_{ji}(\mathbf{x}^*)^{\mathbf{y}_i} = [\mathbf{A}_k]_{ij}(\mathbf{x}^*)^{\mathbf{y}_j} \text{ for each } i \neq j. \tag{41}$$

Finally, to find a **complex balanced realization**, supposed we have the numerical value of one stationary concentration $\mathbf{x}^*$ and consider [43]:

$$\mathbf{Y} \cdot \mathbf{A}_k = \mathbf{Z} \tag{42}$$

$$[\mathbf{A}_k]_{ij} \geq 0 \text{ for } i \neq j \tag{43}$$

$$\mathbf{1}^T \cdot \mathbf{A}_k = \mathbf{0} \tag{44}$$

$$-[\mathbf{A}_k]_{ii}(\mathbf{x}^*)^{\mathbf{y}_i} = \sum_{j \neq i}[\mathbf{A}_k]_{ij}(\mathbf{x}^*)^{\mathbf{y}_j} \text{ for each } i. \tag{45}$$

5.2 Modifying network while preserving dynamical equivalence

In this section, we introduce several modifications to an existing realization while preserving dynamical equivalence. These include:

1. positive or convex combinations of realizations,

2. adding new reactions, and

3. eliminating complexes.

This list is non-exhaustive, but we hope that some of the ideas presented in this section will motivate other types of modifications that preserve dynamical equivalences.

5.2.1 Positive and convex combinations of realizations

Suppose we have found several realizations $(\mathbf{Y}, \mathbf{A}_k^1), (\mathbf{Y}, \mathbf{A}_k^2), \ldots, (\mathbf{Y}, \mathbf{A}_k^I)$ to the kinetic differential equation $\dot{\mathbf{x}} = \mathbf{Z} \cdot \mathbf{x}^{\mathbf{Y}}$. Then we can easily produce other realizations which may be simpler, or fulfil some further properties [16].

Theorem 5. Suppose that $\mathbf{A}_k^1, \ldots, \mathbf{A}_k^I$ are realizations of a kinetic differential equation, with complex set $\mathbf{Y}$, and let $\mathbf{A}_k := \sum_{i=1}^I \alpha_i \mathbf{A}_k^i$ for some $\alpha_i > 0$, $i = 1, \ldots, I$.

1. $\mathbf{A}_k$ is a Kirchhoff matrix, and realizes the equation $\dot{\mathbf{x}} = (\sum_i \alpha_i \mathbf{Y} \cdot \mathbf{A}_k^i)\mathbf{x}^{\mathbf{Y}}$.

2. If $\sum_{i=1}^I \alpha_i = 1$, then $(\mathbf{Y}, \mathbf{A}_k)$ realizes the same kinetic differential equation $\dot{\mathbf{x}} = \mathbf{Z} \cdot \mathbf{x}^{\mathbf{Y}}$.

3. If for all $i = 1, \ldots, I$, the matrix $\mathbf{A}_k^i$ defines a reversible realization of some kinetic equation, then $\mathbf{A}_k$ defines a reversible realization of some kinetic equation.

4. If for all $i = 1, \ldots, I$, the matrix $\mathbf{A}_k^i$ defines a weakly reversible realization of some kinetic equation, then $\mathbf{A}_k$ defines a weakly reversible realization of some kinetic equation.

Proof.

1. Obviously, the off-diagonal terms of $\mathbf{A}_k$ are nonnegative. Also the column sums of $\mathbf{A}_k$ are conserved. Thus $\mathbf{A}_k$ is Kirchoff. We see that

$$\mathbf{Y} \cdot \mathbf{A}_k \cdot \mathbf{x}^{\mathbf{Y}} = \sum_{i=1}^I \alpha_i \mathbf{Y} \cdot \mathbf{A}_k^i \cdot \mathbf{x}^{\mathbf{Y}}.$$

2. This trivially follows from the calculation above.

3. For any $p \neq q$, if $[\mathbf{A}_k]_{pq} > 0$, then $[\mathbf{A}_k^i]_{pq} > 0$ for some $1 \leq i \leq I$. Since $\mathbf{A}_k^i$ defines a reversible realization, it must be the case that $[\mathbf{A}_k^i]_{qp} > 0$, hence $[\mathbf{A}_k]_{qp} > 0$, i.e., $\mathbf{A}_k$ defines a reversible realization.

4. The proof is similar to the above. For any $p \neq q$, if $[\mathbf{A}_k]_{pq} > 0$, there must be some $[\mathbf{A}_k^i]_{pq} > 0$. Since $\mathbf{A}_k^i$ defines a weakly reversible realization, the reaction $q \rightarrow p$ must be part of a cycle, i.e., the off-diagonal terms in $\mathbf{A}_k^i$ corresponding to the edges in this cycle are positive, hence those positions in $\mathbf{A}_k$ are also positive. In other words, this particular cycle appears in the graph defined by $\mathbf{A}_k$ as well.

$\square$

5.2.2   Adding new reactions to a realization

Now we look at modifying an existing realization by adding new reactions. Suppose that $\mathbf{A}_k$ defines a realization with complexes $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N$, and suppose that $\mathbf{y}$ is a convex combination of the remaining complexes, i.e., $\mathbf{y} = \sum_{n=1}^{N} v_n \mathbf{y}_n$ where $v_n \geq 0$ for $n = 1, 2, \ldots, N$ and $\sum_{n=1}^{N} v_n = 1$.

Then we can add the following reactions: for every n such that $v_n \neq 0$, add a reaction from $\mathbf{y}$ to the complex $\mathbf{y}_n$, with rate coefficient v_n. These reactions together contributes $\mathbf{0}$ to the differential equation.

We wonder if we can add reactions in the opposite direction, i.e., reactions of the form $\mathbf{y}_n \rightarrow \mathbf{y}$. There are cases when it is possible and cases when it is not. While we do not have an answer to this question, we illustrate the complexity using two examples.

Example 3. First we show an example where adding a complex situated in the convex cone of the original complexes seems to be superfluous: its effect on the induced kinetic differential equation can be expressed only using reactions among the original complexes. Consider the system

$$
\begin{array}{ccc}
2Y & \rightleftharpoons & 2X + 2Y \\
\Updownarrow & & \Updownarrow \\
0 & \rightleftharpoons & 2X
\end{array}
$$

To add the reaction step $2X \rightarrow X + Y$ to get the network

$$
\begin{array}{ccc}
2Y & \rightleftharpoons & 2X + 2Y \\
\Updownarrow & X + Y & \Updownarrow \\
0 & \rightleftharpoons & 2X
\end{array}
$$
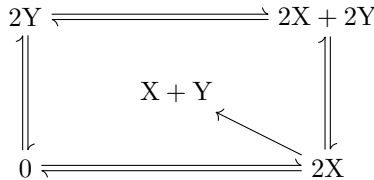
one must borrow from the reactions $2X \rightarrow 0$ and $2X \rightarrow 2X + 2Y$ in the sense that

$$
\underbrace{k_1 x^2 \begin{pmatrix} -2 \\ 0 \end{pmatrix} + k_2 x^2 \begin{pmatrix} 0 \\ 2 \end{pmatrix}}_{\text{in first network}} = \underbrace{(k_1 - \varepsilon) x^2 \begin{pmatrix} -2 \\ 0 \end{pmatrix} + (k_2 - \varepsilon) x^2 \begin{pmatrix} 0 \\ 2 \end{pmatrix} + 2\varepsilon x^2 \begin{pmatrix} -1 \\ 1 \end{pmatrix}}_{\text{in second network}}
$$

where $0 < \varepsilon < \min\{k_1, k_2\}$.

The example above may lead us to believe that it is always possible to add a reaction from some $\mathbf{y}_n$ on the boundary of the convex hall to some complex $\mathbf{y}$ in the interior of the convex hull of other complexes. The example below illustrates that this is not always the case, and further work is needed to explore when it is possible to add such a reaction.

Example 4. If one takes the system

$$2Y \rightleftharpoons 2X + 2Y$$

$$0 \rightleftharpoons 2X$$
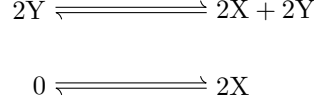
and tries to add the reaction step $2X \to X + Y$ (note that the new complex $X + Y$ is a convex combination of the old ones), there is no way of choosing the rate coefficients so as to get the induced kinetic differential equation of the extended one.

5.2.3 Eliminatin complexes from a realization

Perhaps instead of adding or removing reactions from an existing realization, one may want to add or remove complexes from an existing realization. As noted in [41], it is always possible to add a complex that has not appeared in the current realization, as long as the weighted sum of reaction vectors is $\mathbf{0}$. For example, one can always add the reactions $0 \xleftarrow{k} X \xrightarrow{k} 2X$ without changing the induced kinetic differential equation. The question lies, then, on when it is possible to eliminate a complex from your realization.

Theorem 6. Suppose that $(\mathbf{Y}, \mathbf{A}_k)$ is a realization with $N+1$ complexes $\mathbf{y}_1, \ldots, \mathbf{y}_N, \mathbf{y}_{N+1}$, where the complex $\mathbf{y}_{N+1}$ participates in at least one reaction as a reactant complex. Suppose that the monomial $\mathbf{x}^{\mathbf{y}_{N+1}}$ does not appear explicitly in the kinetic differential equation. Then the differential equation can also be realized using only the complexes $\mathbf{y}_1, \ldots, \mathbf{y}_N$.

Proof. We can write the complex matrix and Kirchoff matrix as

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}' & \mathbf{y}_{N+1} \end{bmatrix}, \quad \text{and} \quad \mathbf{A}_k = \begin{bmatrix} \mathbf{A}'_k & \mathbf{v} \end{bmatrix},$$

where $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N_{\geq 0}$. Then the kinetic differential equation can be written as

$$\frac{d\mathbf{x}}{dt} = \mathbf{Y}'\mathbf{A}'_k\mathbf{x}^{\mathbf{Y}'} + \mathbf{Y}'\mathbf{v}\mathbf{x}^{\mathbf{y}_{N+1}} + \mathbf{y}_{N+1}\mathbf{u}^\top\mathbf{x}^{\mathbf{Y}'} - \left(\sum_{n=1}^N \mathbf{v}_n\right)\mathbf{y}_{N+1}\mathbf{x}^{\mathbf{y}_{N+1}}.$$

Since the complex $\mathbf{y}_{N+1}$ admits some outgoing reactions, we have that $\sum_{n=1}^N \mathbf{v}_n \neq 0$, and, using the fact that the coefficient of $\mathbf{x}^{\mathbf{y}_{N+1}}$ is zero, we find $\mathbf{y}_{N+1} = \frac{1}{\sum_{n=1}^N \mathbf{v}_n}\mathbf{Y}'\mathbf{v}$, and

$$\frac{d\mathbf{x}}{dt} = \mathbf{Y}'\left(\mathbf{A}'_k + \frac{1}{\sum_{n=1}^N \mathbf{v}_n}\mathbf{v}\mathbf{u}^\top\right)\mathbf{x}^{\mathbf{Y}'}.$$

We conclude that $\mathbf{y}_{N+1}$ is a convex combination of the remaining complexes. More importantly, the Kirchoff matrix $\mathbf{A}''_k = \mathbf{A}'_k + \frac{1}{\sum_{n=1}^N \mathbf{v}_n}\mathbf{v}\mathbf{u}^\top$ gives a realization of the same kinetic equation with only the complexes $\mathbf{y}_1, \ldots, \mathbf{y}_N$. □

Remark 2. If $\mathbf{x}^{\mathbf{y}}$ appears explicitly as a term in the kinetic differential equation, then that complex cannot be removed.

After seeing the procedure of eliminating extraneous complexes, one may be tempted to remove all complexes that do not appear as exponents of monomials without ruining desirable properties (e.g., weak reversibility, complex balanced).

Indeed, if one is interested in finding a reversible, or weakly reversible, or detailed balanced, or complex balanced realization, any complex that does not appear in the differential equation (after simplification) can be eliminated from the network [11]:

Theorem 7. A kinetic differential equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ has a complex balanced realization if and only if it has a complex balanced realization where the complexes (the columns of $\mathbf{Y}$) are precisely the monomials in $\mathbf{f}(\mathbf{x})$. Analogous results hold for detailed balanced, weakly reversible, and reversible realizations.

As a consequence, the algorithms described in Section 5.1 can be used to determine whether a given system of kinetic differential equations admits a complex balanced (or detailed balanced, weakly reversible, or reversible) realization. This is a finite calculation as there is no need to compute using different sets of complexes.

6 Relation between weakly reversible and complex balanced realizations for symbolic kinetic equations

When we discussed the algorithm for computing weakly reversible realizations, the reader may have noticed that we introduced an additional Kirchoff matrix $\mathbf{A}'_k$ that does not appear in the realization $(\mathbf{Y}, \mathbf{A}_k)$. This reflects the fact that there is a subtle algebraic relation between weak reversibility and complex balancing. In this section, we explore this relationship.

In order to study the possible network structures that can arise from kinetic differential equation to those with unspecified coefficient, we make the following definition:

Definition 9. For a matrix $\mathbf{Y} \in \mathbb{Z}_{\geq 0}^{M \times N}$ we say that a sign matrix $\mathbf{Z}^{\sigma} \in \{0, +, -\}^{M \times N}$ is **kinetically compatible with** $\mathbf{Y}$ if $\mathbf{Y}_{mr} = 0$ implies $\mathbf{Z}^{\sigma}_{mr} \in \{0, +\}$. In such a case, let

$$\mathcal{Z}(\mathbf{Y}, \mathbf{Z}^{\sigma}) := \left\{ \mathbf{Z} \in \mathbb{R}^{M \times N} : \mathrm{sign}(\mathbf{Z}) = \mathbf{Z}^{\sigma} \right\}, \tag{46}$$

where $\mathrm{sign}(\cdot)$ is the componentwise sign function.

For any $\mathbf{Y} \in \mathbb{Z}_{\geq 0}^{M \times N}$ and kinetically compatible $\mathbf{Z}^{\sigma}$, the set $\mathcal{Z}(\mathbf{Y}, \mathbf{Z}^{\sigma})$ is an unbounded convex cone in $\mathbb{R}^{M \times N}$. According to [24], the differential equation $\dot{\mathbf{x}} = \mathbf{Z} \cdot \mathbf{x}^{\mathbf{Y}}$ is kinetic if $\mathbf{Z} \in \mathcal{Z}(\mathbf{Y}, \mathbf{Z}^{\sigma})$ for any kinetically compatible $\mathbf{Z}^{\sigma}$.

Example 5. The Lotka-Volterra equations

$$\dot{x} = \alpha x - \beta x y$$
$$\dot{y} = \delta x y - \gamma y$$

(with $\alpha, \beta, \gamma, \delta > 0$) can be written in this framework. Let

$$\mathbf{Y} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{Z}^{\sigma} = \begin{pmatrix} + & - & 0 \\ 0 & + & - \end{pmatrix},$$

where clearly $\mathbf{Z}^{\sigma}$ is kinetically compatible with $\mathbf{Y}$. The matrix corresponding to the Lotka-Volterra equations is $\mathbf{Z} = \begin{pmatrix} \alpha & -\beta & 0 \\ 0 & \delta & -\gamma \end{pmatrix}$, which belongs to the set $\mathcal{Z}(\mathbf{Y}, \mathbf{Z}^{\sigma})$.

We defined $\mathcal{Z}(\mathbf{Y}, \mathbf{Z}^{\sigma})$ in order to study possible behaviours a kinetic differential equation may exhibit for some choice of rate coefficients.

Definition 10. Let $\mathbf{Z}^{\sigma}$ be kinetically compatible with the matrix $\mathbf{Y}$. We say that $\mathcal{Z}(\mathbf{Y}, \mathbf{Z}^{\sigma})$ has the **capacity to be complex balanced** (or **capacity to be weakly reversible**) if there exists $\mathbf{Z} \in \mathcal{Z}(\mathbf{Y}, \mathbf{Z}^{\sigma})$ such that $\dot{\mathbf{x}} = \mathbf{Z} \cdot \mathbf{x}^{\mathbf{Y}}$ has a complex balanced (or weakly reversible) realization.

Theorem 8. Let $\mathbf{Z}^{\sigma}$ be kinetically compatible with the matrix $\mathbf{Y}$. The set $\mathcal{Z}(\mathbf{Y}, \mathbf{Z}^{\sigma})$ has the capacity to be complex balanced if and only if it has the capacity to be weakly reversible.

Proof. It is known that complex balancing implies weak reversibility. Suppose there is some $\mathbf{Z} \in \mathcal{Z}(\mathbf{Y}, \mathbf{Z}^\sigma)$ with a weakly reversible realization $(\mathbf{Y}, \mathbf{A}_k)$. Then there is a positive vector $\mathbf{p}$ in the kernel of the Kirchoff matrix $\mathbf{A}_k$. Fix an arbitrary positive state $\mathbf{x}^*$. Consider the following scaled Kirchoff matrix

$$\mathbf{A}'_k = \mathbf{A}_k \cdot diag\left(\frac{\mathbf{p}}{(\mathbf{x}^*)^{\mathbf{Y}}}\right) \tag{47}$$

where the division $\frac{\mathbf{p}}{(\mathbf{x}^*)^{\mathbf{Y}}}$ is componentwise. Then

$$\mathbf{A}'_k(\mathbf{x}^*)^{\mathbf{Y}} = \mathbf{A}_k \cdot diag\left(\frac{\mathbf{p}}{(\mathbf{x}^*)^{\mathbf{Y}}}\right) \cdot (\mathbf{x}^*)^{\mathbf{Y}} = \mathbf{A}_k \cdot \mathbf{p} = \mathbf{0},$$

i.e., $\mathbf{A}'_k$ defines a complex balanced realization with stationary concentration $\mathbf{x}^*$. $\qquad\square$

7   On deficiency zero realizations and uniqueness

Recall that the deficiency of a reaction network is the nonnegative integer $\delta = N - L - S$, where N is the number of complexes, L is the number of linkage classes, and S is the dimension of the stoichiometric subspace $\mathcal{S}$. A network with lower deficiency tends to have nicer dynamical properties; for example, by the Deficiency Zero Theorem or Deficiency One Theorem. Let us formulate a necessary and sufficient condition on network structure for zero deficiency [21].

We say that the vectors $\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_N \in \mathbb{R}^M$ are ***affinely independent*** if $\mathbf{y}_1 - \mathbf{y}_0, \ldots, \mathbf{y}_N - \mathbf{y}_0$ are linearly independent. Note that in this case $\mathbf{y}_0 - \mathbf{y}_n, \ldots, \mathbf{y}_i - \mathbf{y}_n, \ldots, \mathbf{y}_N - \mathbf{y}_n$ are linearly independent for any $n = 1, 2, \ldots, N$ and $i \neq n$. Also, linearly independent vectors are affinely independent.

Theorem 9. A reaction network is deficiency zero if and only if

1. the complex vectors within each linkage class are affinely independent, and

2. the stoichiometric spaces belonging to the linkage class are linearly independent.

Proof. Consider the ℓ-th linkage class. Let N_ℓ be the number of complexes in this linkage class, and let $\mathcal{S}_\ell$ be the stoichiometric space belonging to this linkage class. Let $\delta_\ell := N_\ell - 1 - \dim \mathcal{S}_\ell$ be the deficiency of the ℓ-th linkage class. Finally, let $\mathcal{S}$ be the full stoichiometric space.

The deficiency of the entire network can be related to the deficiencies of the linkage classes:

$$\delta = N - L - S = \sum_{\ell=1}^{L}(N_\ell - 1 - \dim \mathcal{S}_\ell) + \sum_{\ell=1}^{L} \mathcal{S}_\ell - \dim \mathcal{S}$$

$$= \sum_{\ell=1}^{L} \delta_\ell + \sum_{\ell=1}^{L} \dim \mathcal{S}_\ell - \dim \mathcal{S} \geq \sum_{\ell=1}^{L} \delta_\ell.$$

Equality holds if and only if the stoichiometric spaces belonging to the linkage classes $\{\mathcal{S}_\ell\}_{\ell=1}^{L}$ are linearly independent. Moreover, $\delta = 0$ if and only if all $\delta_\ell = 0$. Finally, $\delta_\ell = 0$ implies $\dim \mathcal{S}_\ell = N_\ell - 1$, i.e. the complexes in any given linkage classes are affinely independent. $\qquad\square$

If we restrict ourselves to weakly reversible realizations, the deficiency δ is minimized if the set of complexes are precisely the exponents of monomials [11]. Each additional complex that does not appear as exponents of monomials increases the deficiency δ. It has been shown that any deficiency zero realization with one strong terminal class is unique if the set of complexes is fixed [16]. Combining these two facts, we have the following Proposition.

Proposition 1. Any weakly reversible deficiency zero realization with a single linkage class is unique.

The hypotheses of weakly reversible and deficiency zero in Proposition 1 are required as the following examples illustrate.

Example 6. The reaction networks $Y \xleftarrow{k} 0 \xrightarrow{k} X$ and $0 \xrightarrow{k} X + Y$ are two different deficiency zero realizations with one linkage class, although they are not weakly reversible.

Example 7. The reaction network $0 \rightleftharpoons X \rightleftharpoons 2X$ (with any rate coefficients) and the reaction network
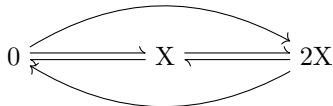
(for wisely chosen rate coefficients) are dynamically equivalent. Both are weakly reversible and with one linkage class, but $\delta = 1$.

Example 8. Two dynamically equivalent weakly reversible realizations where one has $\delta = 1$ are

and
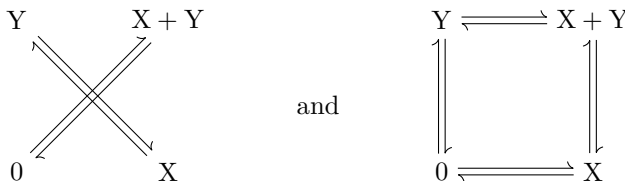
with all rate coefficients are taken to be 1.

Nonetheless, Proposition 1 naturally leads us to the following conjecture, for which a proof has been proposed [10]:

Conjecture 2. Any weakly reversible deficiency zero reaction network has a unique realization.

Throughout this work, we have seen over and over again that realizability (more precisely, identifying the network structure) is an underdetermined problem. Is it possible to use this flexibility in network structure to get more out information about a kinetic differential equation?

An idea comes to mind; we will formulate a conjecture following an example.

Example 9. Consider the mass action system
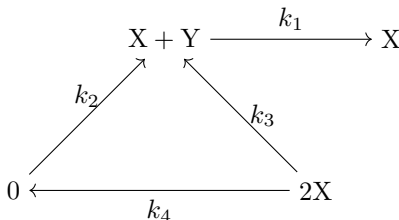
for any rate coefficients k_1, k_2, k_3, $k_4 > 0$. This system has a weakly reversible, deficiency zero realization:
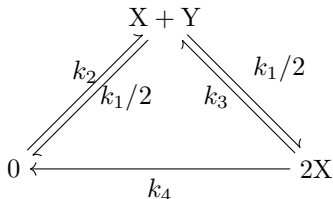
The induced kinetic differential equation of both systems is

$$\dot{x} = k_2 - x^2(k_3 + 2k_4), \qquad \dot{y} = k_2 - k_1 xy + k_3 x^2.$$

In particular, the latter mass action system is complex balanced for all choice of rate coefficients. This implies that the kinetic differential equation enjoys all the dynamical properties of complex balanced systems, although we could have been looking at the network that is not weakly reversible.

Conjecture 3. Consider a reaction network $\mathcal{N}_1$, and assume that there exists a weakly reversible reaction network $\mathcal{N}_2$ of deficiency δ such that any differential equation induced by $\mathcal{N}_1$ and *positive* **k** values can also be induced by $\mathcal{N}_2$ and positive $\widehat{\mathbf{k}}$ values, and viceversa. Then, the set of rate coefficient values of $\mathcal{N}_1$ for which $\mathcal{N}_1$ induces systems that have complex balanced realizations contains an algebraic variety of codimension not larger than δ.

8 Discussion

We have been talking about finding realizations (with certain network properties) of kinetic differential equations. Starting with a system of polynomial differential equations (possibly with symbolic parameters), one could try to find a (Feinberg-Horn-Jackson) reaction graph that induces the differential equation under mass action kinetics. After fixing a set of complexes, algorithms exist to find realizations (Section 5.1). Furthermore, conditions on the resulting reaction graph (e.g., weakly reversible, complex balanced, omit certain reactions) can be incorporated into the algorithms.[4]

We also considered graph operations on the reaction network that will preserve the kinetic differential equation. For example, in certain cases reactions can be added, and complexes can be removed. Our discussion in Section 5.2 is far from complete; it would be interesting to further explore the allowed operations while preserving dynamical equivalence.

In Section 6, we explored the relationship between weakly reversible and complex balanced realizations. Hidden inside the proof of the main theorem in that section is the scaling equation (47). We believe there is information hidden about the manifold in parameter space that decides whether or not a kinetic differential equation can be complex balanced.

Finally, in our discussion of deficiency zero realizations and uniqueness, we posed two conjectures. The first is that weakly reversible deficiency zero networks have unique realization. The latter is more complicated; it asks for what rate coefficients is the reaction network *dynamically equivalent to complex balanced*.

As the reader can see, a theoretical understanding of the relationships between network structure, the sign structure of the differential equation, and dynamical property is lacking. We hope to see further studies exploring the realizability problem.

Acknowledgement

---

[4]Codes for solving the problems in the paper can be requested from the authors.

References

[1] B. Ács, G. Szederknyi, Z.A. Tuza, and Zs. Tuza, *Computing linearly conjugate weakly reversible kinetic structures using optimization and graph theory*, MATCH Commun. Math. Comput. Chem. **74** (2015), 489–512.

[2] G. Ács, G. Szlobodnyik, and G. Szederknyi, *A computational approach to the structural analysis of uncertain kinetic systems*, Computer Physics Communications **228** (2018), 83–95.

[3] D.F. Anderson, *A proof of the global attractor conjecture in the single linkage class case*, SIAM Journal on Applied Mathematics **71** (2011), no. 4, 1487–1508.

[4] C.P.P. Arceo, E.C. Jose, A. Marin-Sanguino, and E.R. Mendoza, *Chemical reaction network approaches to biochemical systems theory*, Mathematical biosciences **269** (2015), 135–152.

[5] R. Aris, *Prolegomena to the rational analysis of systems of chemical reactions*, Archive for Rational Mechanics and Analysis **19** (1965Jan), no. 2, 81–99.

[6] _____, *Mathematical aspects of chemical reaction*, I&EC Fundamentals **61** (1969), no. 6, 17–29.

[7] B. Boros, *On the existence of positive steady states for weakly reversible mass-action systems*, SIAM Journal on Mathematical Analysis **51** (2019), no. 1, 435–449.

[8] L. Cardelli, M. Tribastone, and M. Tschaikowski, *From electric circuits to chemical networks*, arXiv:1812.03308 (2018).

[9] G. Craciun, *Toric differential inclusions and a proof of the global attractor conjecture*, arXiv:1501.02860 (2016).

[10] G. Craciun, J. Jin, and P.Y. Yu, *Uniqueness of kinetic realizations for weakly reversible deficiency zero networks*. To be submitted.

[11] _____, *An efficient characterization of complex-balanced, detailed-balanced, and weakly reversible systems*, arXiv:1812.06214 (2018).

[12] G. Craciun, F. Nazarov, and C. Pantea, *Persistence and permanence of mass-action and power-law dynamical systems*, SIAM Journal on Applied Mathematics **73** (2013), no. 1, 305–329.

[13] G. Craciun and C. Pantea, *Identifiability of chemical reaction networks*, Journal of Mathematical Chemistry **44** (2008), 244–259.

[14] G. Craciun and P.Y. Yu, *Mathematical analysis of chemical reaction systems*, Israel Journal of Chemistry **50** (2018).

[15] D. Csercsik, G. Szederkényi, and K.M. Hangos, *Parametric uniqueness of deficiency zero reaction networks*, J. Math. Chem. **50** (2012), no. 1, 1–8.

[16] D. Csercsik, G. Szederknyi, and K.M. Hangos, *Parametric uniqueness of deficiency zero reaction networks*, Journal of Mathematical Chemistry **50** (2012), 1–8.

[17] M. Dukarić, H. Errami, R. Jerala, T. Lebar, V.G. Romanovski, and A. Tóth J. Weber, *On three genetic repressilator topologies*, Reaction Kinetics, Mechanisms and Catalysis **126** (2019), no. 1, 3–30.

[18] P. Érdi and J. Tóth, *Mathematical models of chemical reactions. Theory and applications of deterministic and stochastic models*, Princeton University Press, Princeton, New Jersey, 1989.

[19] M. Feinberg, *Complex balancing in general kinetic systems*, Arch. Ratl. Mech. Anal. **49** (1972), 187–194.

[20] _____, *Necessary and sufficient conditions for detailed balancing in mass action systems of arbitrary complexity*, Chem. Eng. Sci. **44** (1989), no. 9, 1819–1827.

[21] _____, *Foundations of Chemical Reaction Network Theory*, Springer International Publishing, New York, 2019.

[22] M. Feinberg and F.J.M. Horn, *Chemical mechanism structure and the coincidence of the stoichiometric and kinetic subspaces*, Archive for Rational Mechanics and Analysis **66** (1977), no. 1, 83–97.

[23] M. Gopalkrshnan, E. Miller, and A. Shiu, *A geometric approach to the global attractor conjecture*, SIAM Journal on Applied Dynamical Systems **13** (2014), no. 2, 758–797.

[24] V. Hárs and J. Tóth, *On the inverse problem of reaction kinetics*, Colloquia mathematica societatis jános bolyai, 1979, pp. 363–379.

[25] F. Horn, *Necessary and sufficient conditions for complex balancing in chemical kinetics*, Arch. Ratl. Mech. Anal. **49** (1972), 172–186.

[26] F. Horn and R Jackson, *General mass action kinetics*, Arch. Ratl. Mech. Anal. **47** (1972), 81–116.

[27] M.D. Johnston, D. Siegel, and G. Szederkényi, *Computing weakly reversible linearly conjugate chemical reaction networks with minimal deficiency*, Mathematical Biosciences **241** (2013), 88–98.

[28] M.D. Johnston, D. Siegel, and G. Szederknyi, *Dynamical equivalence and linear conjugacy of chemical reaction networks: new results and methods*, MATCH Commun. Math. Comput. Chem. **68** (2012), 443–468.

[29] _____, *A linear programming approach to weak reversibility and linear conjugacy of chemical reaction networks*, Journal of Mathematical Chemistry **50** (2012), 274–288.

[30] B. Joshi and A. Shiu, *A survey of methods for deciding whether a reaction network is multistationary*, Mathematical Modelling of Natural Phenomena **10** (2015), no. 5, 47–67.

[31] G. Lente, *Deterministic kinetics in chemistry and systems biology: the dynamics of complex reaction networks*, Springer, 2015.

[32] D. Lichtblau, *Symbolic analysis of multiple steady states in a MAPK chemical reaction network*, J. Symb. Comp. **000** (2018), no. 000, 000. submitted.

[33] G. Liptk, G. Szederknyi, and K.M. Hangos, *Computing zero deficiency realizations of kinetic systems*, Systems and Control Letters **81** (2015), 24–30.

[34] V.N. Orlov and L.I. Rozonoer, *The macrodynamics of open systems and the variational principle of the local potential II. Applications*, Journal of the Franklin Institute **318** (1984), no. 5, 315–347.

[35] C. Pantea, *On the persistence and global stability of mass-action systems*, SIAM Journal on Mathematical Analysis **44** (2012), no. 3, 1636–1673.

[36] J. Rudan, G. Szederkényi, and K.M. Hangos, *Efficient computation of alternative structures for large kinetic systems using linear programming*, MATCH Commun. Math. Comput. Chem. **71** (2014), no. 1, 71–92.

[37] J. Rudan, G. Szederkényi, K.M. Hangos, and T. Péni, *Polynomial time algorithms to determine weakly reversible realizations of chemical reaction networks*, J. Math. Chem. (2014), 1–19.

[38] J. Rudan, G. Szederknyi, K. Hangos, and T. Péni, *Polynomial time algorithms to determine weakly reversible realizations of chemical reaction networks*, Journal of Mathematical Chemistry **52** (2014), 1386–1404.

[39] Computing dynamically equivalent realizations of biochemical reaction networks with mass conservation (2013)

[40] S. Schuster and R. Schuster, *Detecting strictly detailed balanced subnetworks in open chemical reaction networks*, J. Math. Chem. **6** (1991), no. 1, 17–40.

[41] G. Szederkényi, *Comment on "identifiability of chemical reaction networks" by G. Craciun and C. Pantea*, Journal of Mathematical Chemistry **45** (2009), no. 4, 1172–1174.

[42] G. Szederkényi, *Computing sparse and dense realizations of reaction kinetic systems*, J. Math. Chem. **47** (2010), 551–568.

[43] G. Szederkényi and K.M. Hangos, *Finding complex balanced and detailed balanced realizations of chemical reaction networks*, J. Math. Chem. **49** (2011), 1163–1179.

[44] G. Szederkényi, K.M. Hangos, and D. Csercsik, *Computing realizations of reaction kinetic networks with given properties*, Coping with complexity: Model reduction and data analysis, 2010, pp. 253–267.

[45] G. Szederkényi, K.M. Hangos, and T. Péni, *Maximal and minimal realizations of reaction kinetic systems: computation and properties*, MATCH Commun. Math. Comput. Chem. **65** (2011), 309–332.

[46] G. Szederkényi, K.M. Hangos, and Z. Tuza, *Finding weakly reversible realizations of chemical reaction networks using optimization*, MATCH Commun. Math. Comput. Chem. **67** (2012), 193–212.

[47] G. Szederkényi, A. Magyar, and K.M. Hangos, *Analysis and control of polynomial dynamic models with biological applications*, Academic Press, London, San Diego, Cambridge, MA, Oxford, 2018.

[48] J. Tóth, *A formális reakciókinetika globális determinisztikus és sztochasztikus modelljéről (On the global deterministic and stochastic models of formal reaction kinetics with applications)*, MTA SZTAKI Tanulmányok **129** (1981), 1–166. In Hungarian.

[49] J. Tóth, A.L. Nagy, and D. Papp, *Reaction Kinetics: Exercises, Programs and Theorems. Mathematica for Deterministic and Stochastic Kinetics*, Springer–Verlag, New York, 2018.

[50] A.I. Volpert and S.I. Hudyaev, *Analyses in classes of discontinuous functions and equations of mathematical physics*, Martinus Nijhoff Publishers, Dordrecht, 1985. Russian original: 1975.

Supplementary

Examples: Realizations with a minimal number of complexes

Simple examples can be treated directly, without detailed application of Theorem 3.

Example 10. If the differential equation $\dot{x} = kx$ is given (no matter what the sign of k is), then we might try to look for an inducing reaction network consisting of the single complex X and find it impossible the minimal reason being that one cannot have a reaction network with less than two complexes. The right hand side contains a single monomial, but the realization should contain more than one complexes.

Example 11. If the differential equation

$$\dot{x} = -k_1 x + 2k_2 y, \quad \dot{y} = 3k_1 x - k_2 y \tag{48}$$

is given with $k_1, k_2 > 0$, then at least four complexes are needed to realize the equation. How do we prove this?

Two complexes are not enough. With the notations of Theorem 3 we have

$$\mathbf{Z} = \begin{pmatrix} -k_1 & 2k_2 \\ 3k_1 & -k_2 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{b} = \mathbf{0} \in \mathbb{R}^2.$$

As $\mathbf{Y}$ is invertible, $\mathbf{v} = \mathbf{Y}^{-1}\mathbf{0} = \mathbf{0}$ should hold. However, $\mathbf{YB} = \mathbf{B} = \mathbf{Z}$ cannot hold, because $\mathbf{Z}$ is not a compartmental matrix.

Three complexes are not enough either. Include $(\begin{array}{cc} \alpha_1 & \alpha_2 \end{array})^\top$ among the complex vectors with unknown nonnegative integers α_1 and α_2 so that this vector is different from the other two we have as yet included. Then our task is to find nonnegative rate coefficients $k_{21}, k_{31}, k_{12}, k_{32}, k_{13}, k_{23}$ (at least two of them positive) so that

$$\begin{pmatrix} 1 & 0 & \alpha_1 \\ 0 & 1 & \alpha_2 \end{pmatrix} \begin{pmatrix} -k_{12}-k_{13} & k_{21} & k_{31} \\ k_{12} & -k_{21}-k_{23} & k_{32} \\ k_{13} & k_{23} & -k_{31}-k_{32} \end{pmatrix} \begin{pmatrix} x \\ y \\ x^{\alpha_1} y^{\alpha_2} \end{pmatrix}$$

$$= \begin{pmatrix} -(k_{12}+k_{13}-k_{13}\alpha_1)x + (k_{21}+k_{23}\alpha_1)y + (k_{31}-(k_{31}+k_{32})\alpha_1)x^{\alpha_1}y^{\alpha_2} \\ (k_{12}+k_{13}\alpha_2)x - (k_{21}+k_{23}-k_{23}\alpha_2)y + (k_{32}-(k_{31}+k_{32})\alpha_2)x^{\alpha_1}y^{\alpha_2} \end{pmatrix}.$$

First of all, one should have zero coefficients before $x^{\alpha_1} y^{\alpha_2}$, thus $\alpha_1 = \frac{k_{31}}{k_{31}+k_{32}}$ $\alpha_2 = \frac{k_{32}}{k_{31}+k_{32}}$, should hold, therefore $1 - \alpha_1 = \alpha_2$ $1 - \alpha_2 = \alpha_1$. Equating the coefficients of x and y here with those of the right hand side gives an equation without solution for $k_{12}, \ldots, \alpha_2$ so that α_1 and α_2 are nonnegative integers, and among the nonnegative numbers k_{nq} at least two are positive. The equations are as follows.

$$k_{12} + k_{13}(1 - \alpha_1) \qquad\qquad = k_{12} + k_{13}\alpha_2 = k_1, \tag{49}$$
$$k_{21} + k_{23}\alpha_1 \qquad\qquad\qquad\qquad = 2k_2, \tag{50}$$
$$k_{12} + k_{13}\alpha_2 \qquad\qquad\qquad\qquad = 3k_1, \tag{51}$$
$$k_{21} + k_{23}(1 - \alpha_2) \qquad\qquad = k_{21} + k_{23}\alpha_1 = k_2. \tag{52}$$

Obviously, (49) and (51) could only hold if $k_1 = 0$, which cannot happen. Similarly, (50) and (52) could only hold if $k_2 = 0$, which cannot happen either.

Four complexes are enough. Finally, one can easily calculate that the complex vectors in the complex matrix

$$\begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 3 & 0 \end{pmatrix}$$

are enough, Eq. (48) is the induced kinetic differential equation of the reaction network

$$\mathrm{X} \xrightarrow{k_1} 3\,\mathrm{Y}, \quad \mathrm{Y} \xrightarrow{k_2} 2\,\mathrm{X}.$$
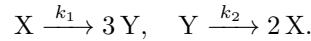
Note that this realization has the minimal number of reactions, (one reaction step is not enough), and is of deficiency zero, as well.

Summing up: here the right hand side of (48) contains two monomials, still there exists no realization with two complexes, more are needed.

Furthermore, three complexes are enough if we allow complexes in $\mathbb{R}^M_{\geq 0}$ instead of $\mathbb{Z}^M_{\geq 0}$. In particular, $\alpha_1 = \frac{4}{5}, \alpha_2 = \frac{3}{5}$ is an appropriate choice. In the pioneering paper [26] this generalization to non-negative real coefficients for complexes is allowed, although here we restrict ourselves to non-negative integer coefficients.

8.1 More examples

Example 12. Invertibility of the complex matrix $\mathbf{Y}$ is not sufficient to have a realization with the exponents of the monomials as complex vectors, as the following example shows.

$$\dot{x} = x^3 - xy^2, \dot{y} = 2x^3$$

implies that

$$\mathbf{Y} = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$$

and the (unique) solution to

$$\begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -b_{21} - u_1 & b_{12} \\ b_{21} & -b_{12} - u_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$$

is

$$b_{12} = -1/3, b_{21} = 1, u_1 = -1, u_2 = 1/3,$$

with some negative numbers. We are forced to formulate a problem: Characterize those matrices with nonnegative (integer) components the inverse of which multiplied with a matrix without negative cross effect gives a compartmental matrix. As this problem again requires finding the feasible solutions of a linear programming (more precisely, MIL) problem, there is not too much hope to find an explicit symbolic solution, but at the same time one can have numerical solution(s) using LP-solvers.

Example 13. Consider the polynomial equation

$$\dot{x} = k_1 x^2 y^5 z - k_{-1} x^3 y^4 z \quad \dot{y} = -k_1 x^2 y^5 z + k_{-1} x^3 y^4 z \quad \dot{z} = 0 \tag{53}$$

with $k_1, k_{-1} > 0$. The canonical representation of this equation is

$$2\,\text{X} + 5\,\text{Y} + \text{Z} \xrightarrow{k_1} 3\,\text{X} + 5\,\text{Y} + \text{Z}, \qquad 3\,\text{X} + 4\,\text{Y} + \text{Z} \xrightarrow{k_{-1}} 2\,\text{X} + 4\,\text{Y} + \text{Z},$$

$$2\,\text{X} + 5\,\text{Y} + \text{Z} \xrightarrow{k_1} 2\,\text{X} + 4\,\text{Y} + \text{Z}, \qquad 3\,\text{X} + 4\,\text{Y} + \text{Z} \xrightarrow{k_{-1}} 3\,\text{X} + 5\,\text{Y} + \text{Z},$$

a weakly reversible reaction network with $N = 4$ complexes, $L = 1$ linkage class and a $S = 2$-dimensional stoichiometric space, thus it has a deficiency $1 = N - L - S$.

Direct calculation shows that the only positive stationary point is

$$x_* = \frac{x_0 + y_0}{1 + K} \quad y_* = \frac{K(x_0 + y_0)}{1 + K} \quad z_* = z_0,$$

however, its existence directly follows from weak reversibility [7].

Example 14. The canonical realization of the differential equation

$$\dot{x} = -4x + 2y + 3z \quad \dot{y} = x - 5y + z + 1 \quad \dot{z} = 2x + 3y - 4z \tag{54}$$

is the irreversible, deficiency three reaction network consisting of a single linkage class of Figure 3. However, it is "obviously" the induced kinetic differential equation of the zero deficiency weakly reversible reaction network in Figure 4. (Here again, the unique positive stationary point can explicitely be determined.)

Let us see a nonlinear example.

Example 15. The canonical realization of the differential equation

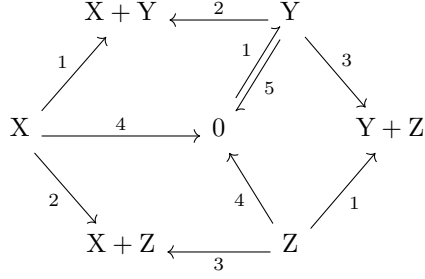$$\dot{x} = -12x^3 + 6y^2, \quad \dot{y} = 2x^3 - 4y^2 + 2 \tag{55}$$

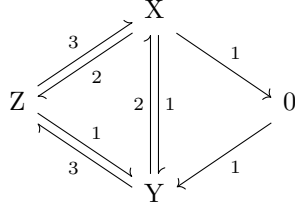Figure 3: The canonical realization of Eq. (54)

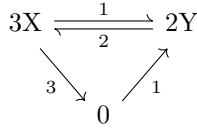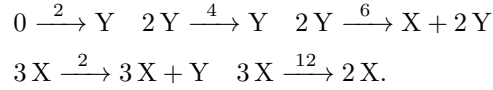Figure 4: A weakly reversible compartmental realization of Eq. (54)

Figure 5: A weakly reversible, zero deficiency reaction network inducing Eq. (55)

is the irreversible, deficiency 3 reaction network consisting of two linkage classes

$$0 \xrightarrow{2} Y \quad 2\,Y \xrightarrow{4} Y \quad 2\,Y \xrightarrow{6} X + 2\,Y$$

$$3\,X \xrightarrow{2} 3\,X + Y \quad 3\,X \xrightarrow{12} 2\,X.$$

However, it is (much less) "obviously" the induced kinetic differential equation of the weakly reversible, zero deficiency reaction network in Figure 5 consisting of a single linkage class. To show this in a formal way, we apply Theorem 3.

With the notations of Theorem 3 we have

$$\mathbf{Z} = \begin{pmatrix} -12 & 6 \\ 2 & -4 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

As $\mathbf{Y}$ is invertible, $\mathbf{v} = \mathbf{Y}^{-1}\mathbf{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ should hold. The unique solution of

$$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -b_{21} - u_1 & b_{12} \\ b_{21} & -b_{12} - u_2 \end{pmatrix} = \begin{pmatrix} -12 & 6 \\ 2 & -4 \end{pmatrix}$$

under the conditions $b_{12}, b_{21}, u_1, u_2 \geq 0$ is calculated with or without the direct use of $\mathbf{Y}^{-1}$. Thus, the unique reaction network to induce Eq. (55) (only having the complexes defined by the monomials on the right hand side) is seen in Figure 5. *How to write the linear equation for the coefficients and why does it tells us that the solution exists and is unique? Here we have four unknowns and four equations, in general we have $N^2 = M^2$ equations and the same number of unknowns.*

Contrary to what the title suggests, here we show a network with thee complexes inducing a differential equation having two terms on the right hand side.

Example 16. The induced kinetic differential equation of the network

$$\alpha_1 X \xrightleftharpoons[k_{-1}]{k_1} \beta_1 X \xrightleftharpoons[k_{-2}]{k_2} \beta_2 X$$
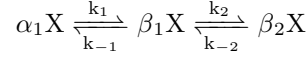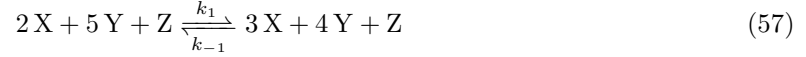
with $\alpha_1 < \beta_1 < \beta_2$ and

$$k_2(\beta_2 - \beta_1) = k_{-1}(\beta_1 - \alpha_1) \tag{56}$$

only has two terms on its right hand side:

$$\dot{x} = k_1 x^{\alpha_1}(\beta_1 - \alpha_1) + k_{-2} x^{\beta_2}(\beta_1 - \beta_2).$$

No wonder, as the condition (56) implies that β_1 is the convex combination of α_1 and β_2.

Example 17. The induced kinetic differential equation of the reaction network

$$2\,X + 5\,Y + Z \xrightleftharpoons[k_{-1}]{k_1} 3\,X + 4\,Y + Z \tag{57}$$

is (53) which we repeat here:

$$\dot{x} = k_1 x^2 y^5 z - k_{-1} x^3 y^4 z \quad \dot{y} = -k_1 x^2 y^5 z + k_{-1} x^3 y^4 z \quad \dot{z} = 0 \tag{58}$$

This is of the form

$$\dot{\mathbf{x}} = k_1 x^2 y^5 z \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + k_{-1} x^3 y^4 z \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \tag{59}$$

therefore an inducing reaction network is just the one in (57).

Mass conserving realizations

The reaction (1) is **_mass conserving_**, if there exists a vector $\varrho \in \mathbb{R}_{>0}^M$ such that for all $r = 1, 2, \ldots, R$

$$\sum_{m=1}^{M} \alpha(m, r)\varrho^m = \sum_{m=1}^{M} \beta(m, r)\varrho^m \tag{60}$$

holds. Note that to ensure mass conservation the existence of a positive vector in the left kernel of the span of the range of the right hand side is necessary but not sufficient, [22, p. 89].

Consider again the differential equation

$$\dot{\mathbf{x}} = \mathbf{Z}\mathbf{x}^{\mathbf{Y}}. \tag{61}$$

According to the definition (60) the existence of a vector $\varrho \in \mathbb{R}_{>0}^M$ such that

$$\varrho^\top \mathbf{Z} = \mathbf{0} \in \mathbb{R}^R \tag{62}$$

is a necessary condition of mass conservation. However, it is only sufficient with some restrictions.

Theorem 10. If (61) has a realization such that the number of linkage classes is equal to the number of ergodic components (terminal strong linkage classes) and there exists a vector $\varrho \in \mathbb{R}_{>0}^M$ such that (62) holds, then this realization is mass conserving.

It is shown in [39] that the same kinetic model may have both mass conserving and non-mass conserving realizations.