The following corrections not affecting the operation of the presented algorithm are made (indicated by blue color in the text):

1. definition of weak reversibility (page 4)

2. numbering of complexes in eq. (28) (page 13)

Finding weakly reversible realizations of chemical reaction networks using optimization

Gábor Szederkényi[a,1], Katalin M. Hangos[a,b], Zsolt Tuza[a,b]

[a] Computer and Automation Research Institute,
Hungarian Academy of Sciences
H-1518, P.O. Box 63, Budapest, Hungary
[b] Faculty of Information Technology, University of Pannonia
H-8200, Egyetem u. 10, Veszprém, Hungary

Abstract

An algorithm is given in this paper for the computation of dynamically equivalent weakly reversible realizations with the maximal number of reactions, for chemical reaction networks (CRNs) with mass action kinetics. The original problem statement can be traced back at least 30 years ago [21]. The algorithm uses standard linear and mixed integer linear programming, and it is based on elementary graph theory and important former results on the dense realizations of CRNs. The proposed method is also capable of determining if no dynamically equivalent weakly reversible structure exists for a given reaction network with a previously fixed complex set.

Keywords: reaction kinetic systems, mass action kinetics, linear programming
AMS classification: 80A30 (chemical kinetics), 90C35 (programming involving graphs or networks)

---

[1]The first author takes all responsibility for any errors or imperfections that might be found in this version of the document.

1 Introduction

Chemical reaction networks (CRNs) are widely used for modeling chemical and biochemical processes in laboratory, industrial or natural environments [38]. Despite their simple mathematical structure, CRN models are capable of describing complex dynamical phenomena such as multiple steady states or oscillating behaviour that can be of key importance in understanding the operation of the modeled technological or living mechanisms [27, 37, 31]. From the point of view of mathematical sciences and systems theory, CRN models and the related analysis methods are also interesting and well-utilizable in the study of the characteristic features of nonlinear dynamical systems [1, 6]. A special class of CRNs is the set of networks whose dynamics is governed by the mass action law (MAL). In this paper, we will deal with such systems.

The foundations of chemical reaction network theory (CRNT) defining the basic notions and properties of reaction networks, were laid in the early papers and lecture notes [20, 14, 15, 16]. Since then, the development in the field has been continuous with significant contributions such as in [21, 13, 30, 18, 8, 7, 9, 29].

It has been known for long that CRNs with different structures and even with different set of chemical complexes may give rise to exactly the same set of differential equations describing the time-evolution of specie concentrations [20]. This fact is called *macro-equivalence* in [20] and it is sometimes referred to as the "fundamental dogma of chemical kinetics". Different CRNs producing the same ODEs will be called *dynamically equivalent* in this paper. Despite of the usefulness of dynamic equivalence in the analysis of CRNs, this property has only been studied to a rather limited extent in the literature. Conditions for the dynamical equivalence to a detailed balanced mechanism were first given in [2]. In [10], the authors gave necessary and sufficient conditions for the dynamic equivalence (also called *confoundability*) of CRNs that was extended with a special case in [32]. In [22], the notion of dynamical equivalence was significantly extended by the introduction of *conjugate networks* that mean CRNs with qualitatively the same dynamics where there is a well-defined mapping that takes trajectories of one system to into trajectories of the other. Also in [22], conditions for conjugacy are given when the above mentioned mapping is a nontrivial linear transformation.

In [33] the terms *dense* and *sparse realizations* were introduced for dynamically equivalent reaction networks containing the maximal or minimal number of reactions with a fixed set of complexes. Additionally, a computational method based on mixed integer linear programming (MILP) was proposed to compute such realizations. In [36], important properties of dense realizations were described: it was shown that the structure (i.e. the unweighted directed graph) of a dense realization of a given CRN is unique, and the unweighted directed reaction graph of any other realization is the subgraph of the dense realization if the set of complexes is fixed. This is a key result that forms the basis of the algorithm presented in this paper. Furthermore, additional MILP based methods were also given in [36] for computing reversible CRN realizations, and realizations with the minimal number of complexes. It was shown in [35] that the computation of dynamically equivalent detailed balanced and complex balanced CRNs can be traced back to simple linear programming (LP).

Many strong results of Chemical Reaction Network Theory apply to weakly reversible networks. The most well-known examples are probably the *Deficiency Zero* and *Deficiency One Theorems* relating complex composition, network structure and the qualitative properties of CRN dynamics [16]. Furthermore, it is known that for any weakly reversible network structure, there always exists

such a parametrization of the reaction rate coefficients that gives a complex balanced system [19]. Finally, the famous *persistency conjecture*, says that the dynamics of weakly reversible networks is persistent in the sense that no trajectory that starts in the positive orthant has an ω-limit point on the boundary of $\mathbb{R}^n_+$.

Similarly to numerous other key CRN properties such as deficiency, full reversibility, complex or detailed balance, weak reversibility is also not the property of the differential equations of the CRN models, but it is a realization property. Among the concluding open questions of [21] (where the authors gave necessary and sufficient conditions for polynomial ODEs to be kinetic together with a constructive proof) we can read the following: "*We may look for a mechanism in a class of mechanisms with a given - chemically relevant - property. Such a property may be conservativity, (weak) reversibility, zero deficiency or just structural stability as well.*" Additionally, in [22] the following is written: "*The development of algorithms and computer software which can efficiently check for viable weakly reversible target networks ... is a primary interest*". Therefore it is of significant importance to work out a method that is able to construct a dynamically equivalent weakly reversible realization for a given CRN if it exists.

Optimization tools based on e.g. linear/nonlinear programming or integer programming have been widely and successfully used in solving scientific and engineering problems [11, 12, 17]. Optimization methods can be very useful tools for deciding the feasibility of complex problems and give feasible solutions (if they exist), even if the original problem is hard to treat analytically [5]. We will follow this approach in this paper.

The structure of the paper is the following. In section 2 the applied models and relevant properties of CRNs will be summarized. Section 3 contains the main contribution, i.e. the algorithm for determining weakly reversible realizations of CRNs. Illustrative examples are shown in section 4, while the conclusions can be found in section 5.

2 Models and properties of chemical reaction networks

2.1 Basic notions for the description of chemical reaction networks

Following [15] and several other works, we will characterize CRNs with the following three sets.

1. $\mathcal{S} = \{X_1, \ldots, X_n\}$ is the set of *species* or chemical substances.

2. $\mathcal{C} = \{C_1, \ldots, C_m\}$ is the set of *complexes*. Formally, the complexes are represented as linear combinations of the species, i.e.

$$C_i = \sum_{j=1}^n \alpha_{ij} X_j, \quad i = 1, \ldots, m \tag{1}$$

   where α_{ij} are nonnegative integers and are called the *stoichiometric coefficients*.

3. $\mathcal{R} = \{(C_i, C_j) \mid C_i, C_j \in \mathcal{C}, \text{ and } C_i \text{ is transformed to } C_j \text{ in the CRN}\}$ is the set of *reactions*. The relation $(C_i, C_j) \in \mathcal{R}$ will be denoted as $C_i \to C_j$. Moreover, a nonnegative weight, the *reaction rate coefficient* denoted by k_{ij} is assigned to each reaction $C_i \to C_j$.

3

The above characterization naturally gives rise to the following graph structure (often called 'Feinberg-Horn-Jackson graph' or simply reaction graph) of a reaction network [16]. The weighted directed graph $G = (V, E)$ of a reaction network consists of a finite nonempty set V of vertices and a finite set E of ordered pairs of distinct vertices called directed edges. The vertices correspond to the complexes, i.e. $V = \{C_1, C_2, \ldots C_m\}$, while the directed edges represent the reactions, i.e. $(C_i, C_j) \in E$ if complex C_i is transformed to C_j in the reaction network. The reaction rate coefficients k_{ij} are assigned as positive weights to the corresponding directed edges $C_i \rightarrow C_j$ in the graph. A set of complexes $\{C_1, \ldots, C_k\}$ is called a *linkage class* of a CRN, if the complexes of the set are linked to each other in the reaction graph but not to any other complex. It is remarked that loops (i.e. directed edges that start and end at the same vertex) are not allowed in reaction graphs. A reaction network is called *reversible*, if whenever it contains the reaction $C_i \rightarrow C_j$, then the reverse reaction $C_j \leftarrow C_i$ is also present in the CRN. A reaction network is called *weakly reversible* if whenever complex C_j is reachable from complex C_i on a directed path in the reaction graph, then C_i is also reachable from C_j on a directed path.

A directed graph is called *strongly connected* if there exists a path from each vertex of the graph to any other vertex. A *strongly connected component* or simply *strong component* of a directed graph is a set of vertices such that the directed edges between them provide a directed path from any vertex of the set to any other vertex, and to which no additional vertex can be added (i.e. a maximal strongly connected subgraph). Since a vertex is naturally reachable from itself through an empty path, strong components containing only one vertex will be called *trivial strong components*. Clearly, the vertices of the individual linkage classes of a weakly reversible CRN form the strong components of the reaction graph.

Assuming mass-action kinetics, the following dynamical description will be used to describe the time-evolution of specie concentrations [15, 16]:

$$\dot{x} = Y \cdot A_k \cdot \psi(x) \tag{2}$$

where x_i denotes the concentration of specie X_i. The ith column of Y contains the composition of complex C_i, i.e. $[Y]_{ij} = \alpha_{ji}$. The structure and parameters of the reaction graph are stored in the column conservation matrix A_k (also called the *Kirchhoff matrix* of the CRN) as follows

$$[A_k]_{ij} = \begin{cases} -\sum_{l=1, l \neq i}^{m} k_{il}, & \text{if} \quad i = j \\ k_{ji}, & \text{if} \quad i \neq j \end{cases} \tag{3}$$

Finally, $\psi : \mathbb{R}^n \mapsto \mathbb{R}^m$ is a monomial-type vector mapping defined by

$$\psi_j(x) = \prod_{i=1}^{n} x_i^{[Y]_{ij}}, \quad j = 1, \ldots, m \tag{4}$$

2.2 Dynamically equivalent realizations

As it has been mentioned in the introduction, reaction networks with different structures and/or parametrizations can give rise to the same kinetic differential equations. Therefore, we will call two reaction networks given by the matrix pairs $(Y^{(1)}, A_k^{(1)})$ and $(Y^{(2)}, A_k^{(2)})$ *dynamically equivalent*, if

$$Y^{(1)} A_k^{(1)} \psi^{(1)}(x) = Y^{(2)} A_k^{(2)} \psi^{(2)}(x) = f(x), \quad \forall x \in \bar{\mathbb{R}}_+^n \tag{5}$$

where for $i = 1, 2$, $Y^{(i)} \in \mathbb{R}^{n \times m_i}$ have nonnegative integer entries, $A_k^{(i)}$ are valid Kirchhoff matrices, and

$$\psi_j^{(i)}(x) = \prod_{k=1}^{m_i} x_k^{[Y^{(i)}]_{kj}}, \quad i = 1, 2, \ j = 1, \ldots, m. \tag{6}$$

In this case, $(Y^{(i)} A_k^{(i)})$ for $i = 1, 2$ are called *realization*s of a kinetic vector field f (see, e.g. [21, 35] for more details). It is also appropriate to call $(Y^{(1)}, A_k^{(1)})$ a *realization* of $(Y^{(2)}, A_k^{(2)})$ and vica versa.

3 The algorithm for computing weakly reversible realizations

We recall that a *dense realization* of a CRN contains the maximal number of reactions (i.e. maximal number of nonzero off-diagonal elements in A_k) with a given stoichiometric matrix Y [33]. Based on the fact that with a given Y, all possible reactions are contained in the structurally unique dense realization [36], a straightforward idea is to try to find a dynamically equivalent weakly reversible mechanism starting from this superstructure.

3.1 Basic principle of the algorithm

Very shortly, the underlying principle of the presented algorithm is that it only removes (if possible) from the dense realization

 (i) edges that cannot be parts of any weakly reversible realization,

 (ii) edges the removal of which is necessarily implied by the deletion of edges belonging to set (i).

The correct operation of the algorithm is based on the following two results.

R1 If each strongly connected component of a directed graph G is contracted to a single vertex, the resulting directed graph is a directed acyclic graph [3]. (A directed graph is called *acyclic* if it has no nontrivial strongly connected subgraphs.)

R2 The structure of the dense realization of any CRN is unique, and the directed unweighted graph of any CRN realization is a subgraph of the directed unweighted graph of the dense realization, if the set of complexes is fixed [36].

From **R2** it follows that for obtaining a CRN superstructure including all possible structures, a dense realization must be computed. Directed edges between different strong components must be removed because they cannot lie on a directed cycle in any realization (**R1**). If this is not possible, then there is no weakly reversible realization of the CRN. However, if the deletion is possible, it may imply the removal of additional reactions because of the linear kinetic constraints (see eqs. (9)-(12)). In general, this may impair the weak reversibility of the obtained network. In such a case, a new dense realization must be computed excluding the unnecessary edges identified in the previous step, and the procedure must be repeated until either a weakly reversible realization is found, or the deletion of undesired edges is no longer possible. In the latter case, no weakly reversible realization of the initial CRN exists with the given stoichiometric matrix Y.

3.2 Definition of input data structure and the necessary additional procedures

We will assume that an initial CRN realization is given with the matrices $(Y^{(0)}, A_k^{(0)})$, and $M = Y^{(0)} \cdot A_k^{(0)}$. The constraint set containing directed edges to be eliminated from the current realization is denoted as

$$\mathcal{K} = \{(p_1, q_1), \ldots, (p_s, q_s)\}, \quad s < r. \tag{7}$$

where p_i and q_i denote the indices of the initial and terminal vertices of the ith edge, respectively, and r is the number of reactions in the CRN.

Now, the following simple procedure will be defined for later use.

$$L = \texttt{FindCrossComponentEdges}(A_k^{in}) \tag{8}$$

The input of the procedure is the Kirchhoff matrix A_k^{in} of a CRN, and the output is a set L containing the directed edges linking different strong components of the reaction graph. The strongly connected components of a directed graph can be determined in linear time using e.g. Kosaraju's, Tarjan's or Gabow's algorithm [3, 25]. Moreover, the examined CRN is weakly reversible if and only if it contains at least 2 reactions, and the output L of $\texttt{FindCrossComponentEdges}$ is empty.

In the following part of this subsection, appropriate modifications of the algorithm described in [33] are given, adapted to the current problem.

3.2.1 Constraints corresponding to mass action dynamics

The main characteristics of mass action dynamics are taken into consideration in the form of the following constraints (see also [33] for more details) in both of the procedures presented in subsections 3.2.2 and 3.2.3.

$$Y \cdot A_k = M \tag{9}$$

$$\sum_{i=1}^{m} [A_k]_{ij} = 0, \quad j = 1, \ldots, m \tag{10}$$

$$[A_k]_{ij} \geq 0, \quad i, j = 1, \ldots, m, \quad i \neq j \tag{11}$$

$$[A_k]_{ii} \leq 0, \quad i = 1, \ldots, m \tag{12}$$

where Y and M are known, and the decision variables are the off-diagonal elements of A_k. It's easy to see that constraints (10)-(12) represent the fact that we are searching for a valid Kirchhoff matrix.

3.2.2 Checking whether a set of reactions is removable from a CRN realization

We call a set of reactions $\mathcal{K}$ *removable* from a CRN realization, if there exists a dynamically equivalent CRN realization that does not contain the directed edges in $\mathcal{K}$. To check this, it is worth separately defining the following LP-based and thus polynomial time procedure to avoid unnecessary MILP computations that are known to be NP-hard.

6

The constraints (9)-(12) are completed with the following ones

$$[A_k]_{q_i,p_i} = 0 \ \text{ for } \ i = 1, \ldots, s \tag{13}$$

where $(p_i, q_i) \in \mathcal{K}$ as it is written in eq. (7). The feasibility of the linear constraints (9)-(12) and (13) can be checked by adding e.g. the following linear objective function to be minimized:

$$G_{LP}(A_k) = \sum_{\substack{i,j=1 \\ i \neq j, \ (j,i) \notin \mathcal{K}}}^{m} [A_k]_{ij} \tag{14}$$

Clearly, eqs. (9)-(14) form a standard linear programming problem, the feasibility of which can be checked in polynomial time [11].

Based on the above, we will define the procedure to check whether a set $\mathcal{K}$ of directed edges is removable from a CRN realization or not in the following way:

$$F_{out} = \texttt{IsRemovable}(Y^{(i)}, A_k^{(i)}, \mathcal{K}), \tag{15}$$

where the input data are as follows: $(Y^{(i)}, A_k^{(i)})$ is a CRN realization, and $\mathcal{K}$ is a constraint set of the form (7) containing the tail and head index pairs of the directed edges to be removed from $(Y^{(i)}, A_k^{(i)})$. The output F_{out} is a boolean variable: its value is $\texttt{true}$ if there exists a dynamically equivalent realization to $(Y^{(i)}, A_k^{(i)})$ not containing the edges listed in K, and it is $\texttt{false}$ if there does not exist such realization.

3.2.3 Computing the dense realization excluding given directed edges

For the solution of this subtask, the well-known results on the combination of propositional logic with mixed integer programming are applied [26, 4]. According to these, a propositional logic problem, where a statement must be proved to be true if a set of compound statements are also given, can be solved through a linear integer problem.

If the procedure $\texttt{IsRemovable}$ returns a $\texttt{true}$ value, the dense realization of the CRN can be computed subject to the constraint that the edges listed in $\mathcal{K}$ are excluded from it. To define the corresponding MILP problem, first we add exactly the same linear constraints contained in eqs. (9)-(13) as in the previous case. To make the forthcoming problem computationally tractable, we also introduce the following bounds for the decision variables

$$[A_k]_{ij} \leq u_{ij}, \ \ u_{ij} > 0, \ \ i,j = 1, \ldots, m, \ \ i \neq j, \ \ (j,i) \notin \mathcal{K} \tag{16}$$

$$[A_k]_{ii} \geq l_i, \ \ l_i < 0, \ \ i = 1, \ldots, m \tag{17}$$

Here we are searching for such A_k that contains the maximal number of nonzero off-diagonal elements. For this, logical variables denoted by δ are introduced and the following compound statements are constructed

$$\delta_{ij} = 1 \leftrightarrow [A_k]_{ij} > \epsilon, \ \ i,j = 1, \ldots, m, \ \ i \neq j, \ \ (j,i) \notin \mathcal{K} \tag{18}$$

where the symbol "↔" denotes "if and only if", and $0 < \epsilon \ll 1$ (i.e. elements of A_k below ϵ are treated as zero). Considering also (16), statement (18) can be translated to the following linear inequalities

$$0 \leq [A_k]_{ij} - \epsilon\delta_{ij}, \quad i,j = 1,\ldots,m, \quad i \neq j, \quad (j,i) \notin \mathcal{K} \tag{19}$$
$$0 \leq -[A_k]_{ij} + u_{ij}\delta_{ij}, \quad i,j = 1,\ldots,m, \quad i \neq j \quad (j,i) \notin \mathcal{K} \tag{20}$$

Now it is possible to compute the realization containing the maximal number of reactions by maximizing the objective function

$$G_{MILP}(\delta) = \sum_{\substack{i,j = 1 \\ i \neq j, \ (j,i) \notin \mathcal{K}}}^{m} \delta_{ij}. \tag{21}$$

Finally, the following procedure is defined based on the MILP problem given by eqs. (9)-(13), (16)-(17), and (19)-(21).

$$A_k^d = \texttt{FindConstrDenseRealization}(Y^{(i)}, A_k^{(i)}, \mathcal{K}), \tag{22}$$

where the input data set is the same as in the case of the procedure $\texttt{IsRemovable}$ (see eq. (15) and the corresponding description). The output A_k^d is a dense realization that does not contain the directed edges listed in the set $\mathcal{K}$. If the set $\mathcal{K}$ is empty, then the procedure is the same that computes dense realizations and that was published in [33].

3.3 Formal description of the algorithm

Now we are in the position to give the formal description of the procedure for determining weakly reversible CRN realizations. The input data of the procedure is an initial CRN realization $(Y^{(0)}, A_k^{(0)})$. The output is an $m \times m$ matrix that is the Kirchhoff matrix of the weakly reversible realization if there exists such, or a zero matrix if the procedure found no weakly reversible realizations. In the algorithm pseudocode, the auxiliary variable $ExitCondition$ is a boolean storing the exit condition from the main loop. The complete pseudocode of the procedure called $\texttt{FindWeaklyReversibleRealization}$ with common notations and keywords can be found in Table 1.

3.4 Main properties of the algorithm

The above described algorithm always finds a dynamically equivalent weakly reversible realization, if it exists. This clearly follows from the results **R1**, **R2** and the basic principles of the algorithm described in subsection 3.1. From these facts it also follows that the algorithm finds the 'densest' weakly reversible realization that structurally contains any other weakly reversible realizations (for illustration, see the results in subsection 4.1). An apparent drawback of the algorithm that it contains a step where a MILP problem has to be solved that is known to be NP-hard. However, the parallel (columnwise) implementation of the procedure $\texttt{FindConstrDenseRealization}$ is possible as it is explained in [33], that significantly extends the number of complexes that can be handled in practice. We remark that the immediate deletion of the columns/rows corresponding to the isolated

```
A_k^out=FindWeaklyReversibleRealization(Y^(0),A_k^(0))
```

1	$A_k^{out}:=0 \in \mathbb{R}^{m \times m}$; $ExitCondition$:=**false**;
2	$Y := Y^{(0)}$; $A_k := A_k^{(0)}$; F_{out}:=**true**; $\mathcal{K} := \{\}$; $L := \{\}$;
3	**while** ($ExitCondition$=**false**) **do**
4	**begin**
5	**if** ($\mathcal{K} \neq \{\}$) **then** F_{out}:=`IsRemovable`($Y,A_k,\mathcal{K}$);
6	**if** (F_{out} =**true**) **then**
7	**begin**
8	A_k:=`FindConstrDenseRealization`($Y,A_k,\mathcal{K}$);
9	L:=`FindCrossComponentEdges`(A_k);
10	**if** ($L = \{\}$) **then** $ExitCondition$:=**true**; A_k^{out}:=A_k;
11	**else** $\mathcal{K} := \mathcal{K} \cup L$;
12	**end**
13	**else** $ExitCondition$:=**true**;
14	**end**
15	**return** A_k^{out};

Table 1: Pseudocode of the algorithm for finding weakly reversible realizations

complexes from matrices Y and A_k after calling the procedure `FindConstrDenseRealization` is also possible, but this requires the renumbering of complexes. This is a technical detail of implementation and does not affect the principle or final output of the algorithm. (However, decreasing the number of optimization variables and constraints in each step in such a way might be required especially in the case of larger networks.)

4 Examples

The following examples were implemented in the MATLAB R13 computation environment using the YALMIP and the Multi-Parametric Toolboxes [24, 23]. The examples were run on a desktop PC with dual Intel Xeon 1.8GHz CPU and 2 Gigabytes of RAM. The implementation of dense realization computation was not parallel, therefore all the decision variables and constraints were put into a single optimization problem in procedure `FindConstrDenseRealization`. The strong components of the reaction graphs were identified using Kosaraju's algorithm [28].

4.1 Weakly reversible realizations of a simple irreversible network

The simple network that can be seen in Fig. 1 a) was taken from [22] (Example 3). In [22], it is shown that for any positive ϵ, the network has a possible weakly reversible realization with the structure shown in Fig. 1b). (The computation and analysis of the parameters of the CRN in Fig.
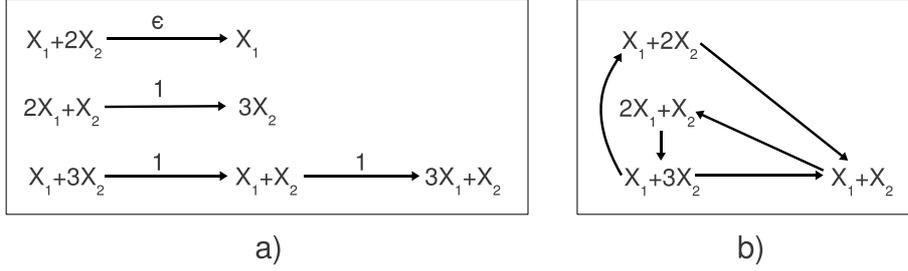
Figure 1: a) Simple irreversible network from [22], b) Structure of one of its possible weakly reversible realizations determined in [22]

1b can be found in [22].) The stoichiometric matrix of the network is

$$Y = \begin{bmatrix} 1 & 1 & 2 & 0 & 1 & 1 & 3 \\ 2 & 0 & 1 & 3 & 3 & 1 & 1 \end{bmatrix} \tag{23}$$

The nonzero elements of the Kirchhoff matrix $A_k \in \mathbb{R}^{7\times7}$ with $\epsilon = 1.5$ are

$$[A_k]_{2,1} = 1.5,\ [A_k]_{4,3} = 1,\ [A_k]_{6,5} = 1,\ [A_k]_{7,6} = 1\ . \tag{24}$$

For this network, the algorithm described in section 3.3 and Table 1 works as follows. After the initialization steps, the dense realization containing all possible reactions with an empty constraint set $\mathcal{K}$ is computed (line 8 of the pseudocode). The Kirchhoff matrix of the dense realization is given by

$$A_k^{(1)} = \begin{bmatrix} -1.25 & 0 & 0.1 & 0 & 0.1 & 0.1 & 0 \\ 0.55 & 0 & 0.1 & 0 & 0.4333 & 0.5 & 0 \\ 0.1 & 0 & -1.4 & 0 & 0.1 & 0.1 & 0 \\ 0.3 & 0 & 0.8 & 0 & 0.3 & 0.1 & 0 \\ 0.1 & 0 & 0.2 & 0 & -1.1333 & 0.1 & 0 \\ 0.1 & 0 & 0.1 & 0 & 0.1 & -1.9 & 0 \\ 0.1 & 0 & 0.1 & 0 & 0.1 & 1 & 0 \end{bmatrix} \tag{25}$$

The structure of the dense realization is shown in Fig. 2. The following steps can be followed using Fig. 3. The dense realization is not weakly reversible because there are edges between different strong components (line 9). The complexes of the single nontrivial strong component are indicated by boldface labels in Fig. 3. It is clear from the figure that the edges adjacent to the complexes X_1, $3X_2$, $3X_1 + X_2$ are to be removed. These edges are drawn with dotted arrows in the figure. Thus, the constraint list is

$$\mathcal{K} = \{(1,2),(1,4),(1,7),(3,2),(3,4),(3,7),(5,2),(5,4),(5,7),(6,2),(6,4),(6,7)\}.$$

The next iteration of the algorithm (line 5) gives that these edges can be removed from the realization. Now a new dense realization is computed excluding edges in $\mathcal{K}$ (line 8). The reactions of this dense realization are indicated by thick arrows in Fig. 3. Note that the constraint of excluding the
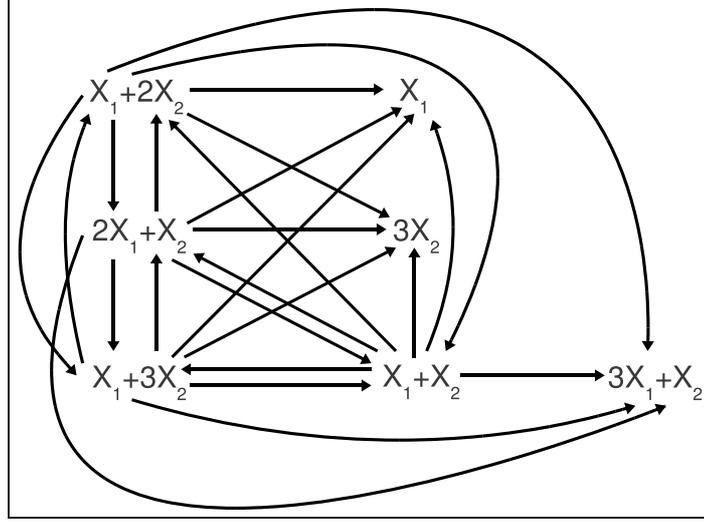
10

Figure 2: Structure of the dense realization of the reaction network shown in Fig. 1 a)

reactions in $\mathcal{K}$ from the network resulted in the removal of directed edges $X_1 + X_2 \rightarrow X_1 + 2X_2$, $X_1 + 2X_2 \rightarrow 2X_1 + X_2$, $X_1 + 3X_2 \rightarrow 2X_1 + X_2$, and $X_1 + X_2 \rightarrow X_1 + 3X_2$, too, that were within the nontrivial strong component of the previous step. In this case, the resulting network remained weakly reversible (lines 9-10), so the algorithm can stop with success and return the determined dynamically equivalent weakly reversible CRN. The structure of the resulting network is shown in Fig. 4. The Kirchhoff matrix of the obtained realization is the following

$$
A_k^{(2)} = \begin{bmatrix}
-3.2 & 0 & 1.8 & 0 & 0.1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -2 & 0 & 0 & 2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.1 & 0 & 0.1 & 0 & -1.05 & 0 & 0 \\
3.1 & 0 & 0.1 & 0 & 0.95 & -2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix} \tag{26}
$$

The running time of the algorithm was 11s using the hardware/software environment described in the beginning of section 4.

It is interesting to note that the obtained weakly reversible realization is not complex balanced. (For convenience, we briefly define the notion of complex balance: A CRN realization (Y, A_k) is called *complex balanced* if there exists a positive vector $x^* \in \mathbb{R}_+^n$ for which $A_k \psi(x^*) = 0$ [20, 30, 7].) However, using the polynomial-time algorithm described in [35], we can compute a complex balanced realization with 6 reactions in 0.1s, that is shown in Fig. 5. It is easy to see that the unweighted directed graphs of Figs. 1 b) and Fig. 5 are the proper subgraphs of the structure visible in Fig. 4.
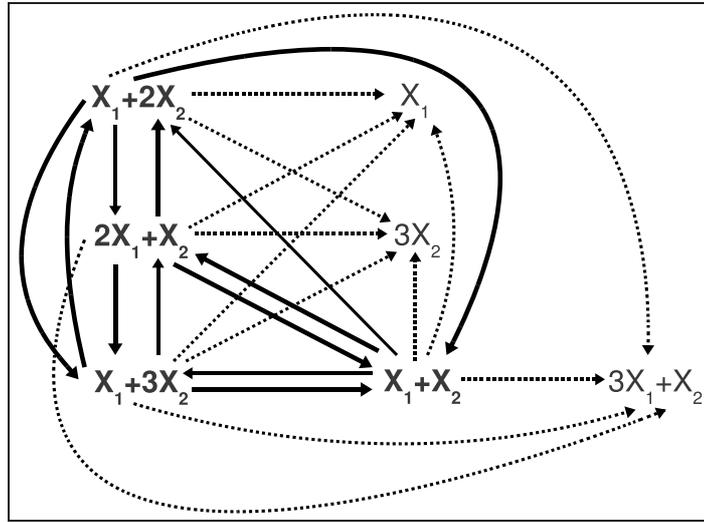
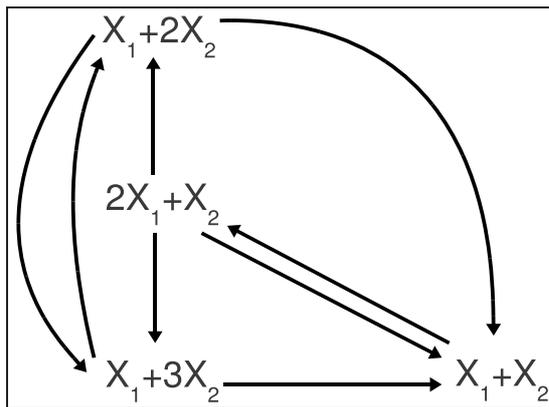Figure 3: Illustration of the operation of the algorithm on the example given in section 4.1

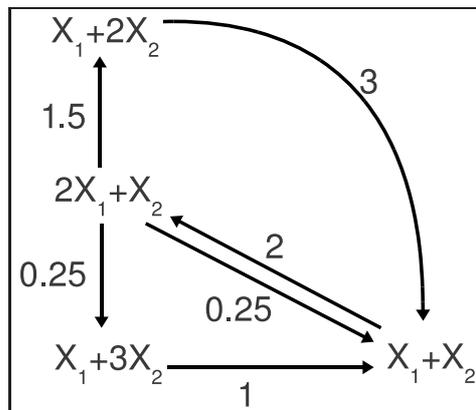Figure 4: The structure of the obtained dynamically equivalent weakly reversible CRN

Figure 5: Complex balanced realization of the CRN described in section 4.1

4.2 Weakly reversible realization of a kinetic polynomial system

The equations and initial CRN in this example were also used in [35] and [34] for illustrating other computation methods. We start from the polynomial ordinary differential equations given by

$$
\begin{aligned}
\dot{x}_1 &= x_3^2 - x_1 x_2 + x_3 x_4 - 2 x_1 x_2^2 x_3 \\
\dot{x}_2 &= x_3^2 - x_1 x_2 + 2 x_3 x_4 - 4 x_1 x_2^2 x_3 \\
\dot{x}_3 &= -2 x_3^2 + x_1 x_2 - x_1 x_2^2 x_3 + 2 x_4^3 \\
\dot{x}_4 &= x_1 x_2 - x_3 x_4 + 4 x_1 x_2^2 x_3 - 3 x_4^3
\end{aligned}
\tag{27}
$$

Using the inverse kinetic algorithm first published in [21] (see also [35] for a summary about the kinetic realizability of polynomial vector fields), a CRN shown in Fig. 6 can be constructed algorithmically that gives the dynamics (27). The numbering of complexes in the figure is the following:

$$
\begin{aligned}
&C_1 = 2X_3,\ C_2 = X_1 + 2X_3,\ C_3 = X_1 + X_2,\ C_4 = X_2,\ C_5 = X_3 + X_4,\ C_6 = X_1 + X_3 + X_4, \\
&C_7 = X_1 + 2X_2 + X_3,\ C_8 = 2X_2 + X_3,\ C_9 = X_2 + 2X_3,\ C_{10} = X_1,\ C_{11} = X_2 + X_3 + X_4, \\
&C_{12} = X_1 + X_2 + X_3,\ C_{13} = X_3,\ C_{14} = X_1 + 2X_2,\ C_{15} = 3X_4, \\
&C_{16} = X_3 + 3X_4,\ C_{17} = X_1 + X_2 + X_4,\ C_{18} = X_1 + 2X_2 + X_3 + X_4,\ C_{19} = 2X_4.
\end{aligned}
\tag{28}
$$

The dense realization of the network contains 80 reactions, therefore it is not shown in a figure. For the sake of completeness, the list of the reactions (i.e. weighted directed edges) of the dense realization in the form (*source vertex number, destination vertex number, rate coefficient*) is given below:

(5, 1, 0.5), (7, 1, 0.1), (11, 1, 0.1), (15, 1, 0.8), (1, 2, 0.1), (3, 2, 0.1), (5, 2, 0.1), (7, 2, 0.1), (12, 2, 0.3), (1, 3, 0.1), (5, 3, 0.1), (7, 3, 0.1), (12, 3, 0.1), (1, 4, 0.1), (3, 4, 0.4), (5, 4, 0.1), (7, 4, 0.1), (11, 4, 0.1), (3, 5, 0.1), (7, 5, 0.1), (11, 5, 0.1), (15, 5, 0.1), (3, 6, 0.1), (5, 6, 0.1), (7, 6, 0.1), (1, 7, 0.1), (3, 7, 0.1), (5, 7, 0.2), (12, 7, 0.3), (1, 8, 0.1), (3, 8, 0.1), (5, 8, 0.3), (7, 8, 0.3), (11, 8, 1.2), (1, 9, 0.1), (5, 9, 0.1), (7, 9, 0.1), (11, 9, 0.2), (1, 10, 0.5), (3, 10, 0.8), (5, 10, 0.1), (7, 10, 0.1), (12, 10, 0.1), (3, 11, 0.1), (5, 11, 0.1), (7, 11, 0.1), (1, 12, 0.1), (3, 12, 0.1), (5, 12, 0.1), (7, 12, 0.1), (1, 13, 0.1), (3, 13, 0.1), (5, 13, 0.1), (7, 13, 0.1), (11, 13, 0.1), (15, 13, 0.1), (1, 14, 0.1), (3, 14, 0.1), (5, 14, 0.1), (7, 14, 0.1), (12, 14, 0.1), (3, 15, 0.1), (5, 15, 0.1), (7, 15, 0.7), (11, 15, 0.1), (5, 16, 0.25), (7, 16, 0.3), (11, 16, 0.7), (15, 16, 0.2), (3, 17, 0.1), (5, 17, 0.1), (7, 17, 0.1), (3, 18, 0.1), (5, 18, 0.1), (7, 18, 0.4), (3, 19, 0.1), (5, 19, 0.1), (7, 19, 0.1), (11, 19, 0.1), (15, 19, 0.1).

The dense realization has 13 strong components. Out of these, there is only one nontrivial strongly connected component containing the vertices 1, 3, 5, 7, 11, 12, and 15. After identifying all directed edges linking different strong components, we obtain that the following 53 edges given in the form (*source vertex number, destination vertex number*) should be deleted from the dense realization in the next step of the algorithm:

(1,2), (3,2), (5,2), (7,2), (12,2), (1,4), (3,4), (5,4), (7,4), (11,4), (3,6), (5,6), (7,6), (1,8), (3,8), (5,8), (7,8), (11,8), (1,9), (5,9), (7,9), (11,9), (1,10), (3,10), (5,10), (7,10), (12,10), (1,13), (3,13), (5,13), (7,13), (11,13), (15,13), (1,14), (3,14), (5,14), (7,14), (12,14), (5,16), (7,16), (11,16), (15,16), (3,17), (5,17), (7,17), (3,18), (5,18), (7,18), (3,19), (5,19), (7,19), (11,19), (15,19).
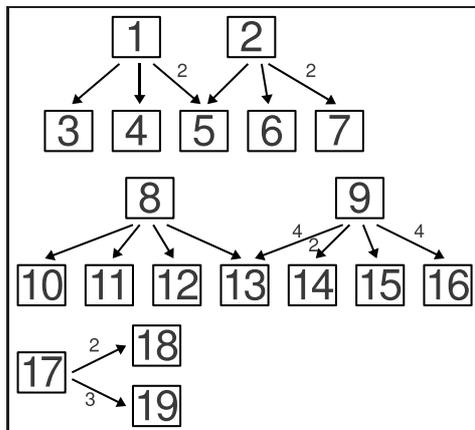
Figure 6: CRN realizing the dynamics of eq. (27). Only reaction rates different from 1 are indicated.

All the above listed edges were possible to remove from the dense realization (their removal implied the deletion of 12 additonal reactions), and the resulting constrained dense realization is shown in Fig.7. It is clear from the figure that edges adjacent to vertices no. 11 and 12 have to be removed in the following step. This final step is illustrated in Fig. 8, where the meaning of the line types is the same as in the case of Fig. 3. With the removal of edges (5,11), (5,12), (7,11), (7,12), the directed edges (5,1), (7,1), (7,3), (5,3), (7,5) and (5,15) were also deleted, and the resulting weakly reversible realization (of deficiency 0) is shown in Fig. 9. The total running time of the algorithm was 80.5s.
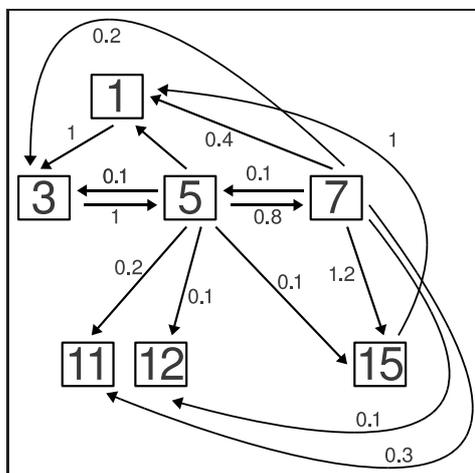

Figure 7: CRN realization of eq. (27) containing the vertices of the only nontrivial strong component of the dense realization
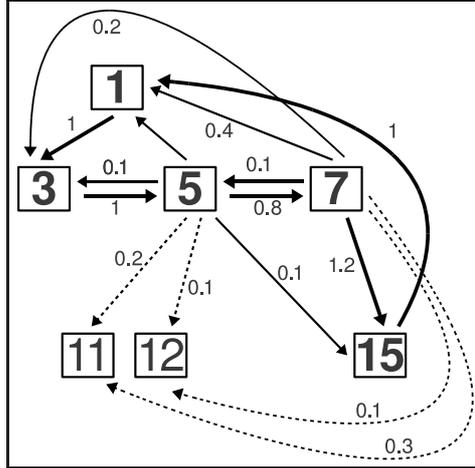
Figure 8: Illustration of the final step of the algorithm on the CRN corresponding to eq. (27)
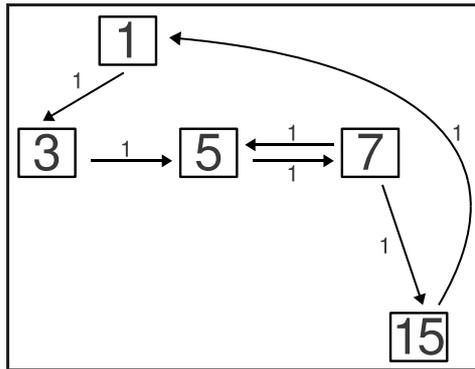

Figure 9: The obtained weakly reversible realization of eq. (27)

5 Conclusions

A numerical algorithm for finding dynamically equivalent weakly reversible realizations of chemical reaction networks was proposed in this paper. The algorithm computes a maximal superstructure that contains all other possible weakly reversible structures as proper subgraphs, and it is able to determine if no weakly reversible realization exists. The method uses linear and mixed integer linear programming steps and it is based primarily on the computation and properties of dense realizations published in [33, 36]. The computationally critical MILP step of the algorithm can be implemented parallelly. The operation of the algorithm has been illustrated on numerical examples.

Acknowledgements

References

[1] D. Angeli. A tutorial on chemical network dynamics. *European Journal of Control*, 15:398–406, 2009.

[2] E. D. Averbukh. Complex balanced kinetic functions in inverse problems of chemical kinetics. *Automation and Remote Control*, 55:1723–1732, 1994.

[3] J. Bang-Jensen and G. Gutin. *Digraphs: Theory, Algorithms and Applications*. Springer, 2001.

[4] A. Bemporad and M. Morari. Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35:407–427, 1999.

[5] S. Boyd and L Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[6] V. Chellaboina, S. P. Bhat, W. M. Haddad, and D. S. Bernstein. Modeling and analysis of mass-action kinetics – nonnegativity, realizability, reducibility, and semistability. *IEEE Control Systems Magazine*, 29:60–78, 2009.

[7] G. Craciun, A. Dickenstein, A. Shiu, and B. Sturmfels. Toric dynamical systems. *Journal of Symbolic Computation*, 44:1551–1565, 2009.

[8] G. Craciun and M. Feinberg. Multiple equilibria in complex chemical reaction networks: I. The injectivity property. *SIAM Journal on Applied Mathematics*, 65 (5):1526–1546, 2005.

[9] G. Craciun and M. Feinberg. Multiple equilibria in complex chemical reaction networks: Semiopen mass action systems. *SIAM Journal on Applied Mathematics*, 70:1859–1877, 2010.

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

[11] G. B. Dantzig and M. N. Thapa. *Linear programming 1: Introduction.* Springer-Verlag, 1997.

[12] G. B. Dantzig and M. N. Thapa. *Linear Programming 2: Theory and Extensions.* Springer-Verlag, 2003.

[13] Gy. Farkas. Kinetic lumping schemes. *Chemical Engineering Science*, 54:3909–3915, 1999.

[14] M. Feinberg. Mathematical aspects of mass action kinetics. In L. Lapidus and N. Amundson, editors, *Chemical Reactor Theory: A Review.* Prentice Hall, Englewood Cliffs, 1977.

[15] M. Feinberg. *Lectures on chemical reaction networks.* Notes of lectures given at the Mathematics Research Center, University of Wisconsin, 1979.

[16] M. Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors - I. the deficiency zero and deficiency one theorems. *Chemical Engineering Science*, 42 (10):2229–2268, 1987.

[17] C.A. Floudas. *Nonlinear and mixed-integer optimization.* Oxford University Press, 1995.

[18] A.N. Gorban, I.V. Karlin, and A.Y. Zinovyev. Invariant grids for reaction kinetics. *Physica A*, 33:106–154, 2004.

[19] F. Horn. Necessary and sufficient conditions for complex balancing in chemical kinetics. *Archive for Rational Mechanics and Analysis*, 49:172–186, 1972.

[20] F. Horn and R. Jackson. General mass action kinetics. *Archive for Rational Mechanics and Analysis*, 47:81–116, 1972.

[21] V. Hárs and J. Tóth. On the inverse problem of reaction kinetics. In M. Farkas and L. Hatvani, editors, *Qualitative Theory of Differential Equations*, volume 30 of *Coll. Math. Soc. J. Bolyai*, pages 363–379. North-Holland, Amsterdam, 1981.

[22] M. D. Johnston and D. Siegel. Linear conjugacy of chemical reaction networks, 2011. Available at arXiv:1101.1663v1 [math.DS].

[23] M. Kvasnica, P. Grieder, and M. Baotić. Multi-Parametric Toolbox (MPT), 2004.

[24] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[25] E. Nuutila and E. Soisalon-Soininen. On finding the strongly connected components in a directed graph. *Information Processing Letters*, 49:9–14, 1994.

[26] R. Raman and I.E. Grossmann. Modelling and computational techniques for logic based integer programming. *Computers and Chemical Engineering*, 18:563–578, 1994.

[27] S. K. Scott. *Oscillation, Waves, and Chaos in Chemical Kinetics.* Oxford University Press, 1994.

[28] M. Sharir. A strong-connectivity algorithm and its application in data flow analysis. *Comput. Math. Appl.*, 7:67–72, 1981.

[29] G. Shinar and M. Feinberg. Structural sources of robustness in biochemical reaction networks. *Science*, 327:1389–1391, 2010.

[30] D. Siegel and D. MacLean. Global stability of complex balanced mechanisms. *Journal of Mathematical Chemistry*, 27:89–110, 2000.

[31] E. Sontag. Structure and stability of certain chemical networks and applications to the kinetic proofreading model of T-cell receptor signal transduction. *IEEE Trans. Autom. Control*, 46:1028–1047, 2001.

[32] G. Szederkényi. Comment on "Identifiability of chemical reaction networks" by G. Craciun and C. Pantea. *Journal of Mathematical Chemistry*, 45:1172–1174, 2009.

[33] G. Szederkényi. Computing sparse and dense realizations of reaction kinetic systems. *Journal of Mathematical Chemistry*, 47:551–568, 2009.

[34] G. Szederkényi. Computing reaction kinetic realizations of positive nonlinear systems using mixed integer programming. In *8th IFAC Symposium on Nonlinear Control Systems - NOLCOS 2010, Bologna, Italy, 1-3 September*, page ThP04.2, 2010.

[35] G. Szederkényi and K. M. Hangos. Finding complex balanced and detailed balanced realizations of chemical reaction networks. *Journal of Mathematical Chemistry*, accepted:in press, doi: 10.1007/s10910–011–9804–9, 2011. Available on the arXiv at http://arxiv.org/abs/1010.4477.

[36] 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(2):309–332, 2011.

[37] R. Thomas and M. Kaufman. Multistationarity, the basis of cell differentiation and memory. I. Structural conditions of multistationarity and other nontrivial behaviour. *Chaos*, 11:170–179, 2001.

[38] P. Érdi and J. Tóth. *Mathematical Models of Chemical Reactions. Theory and Applications of Deterministic and Stochastic Models*. Manchester University Press, Princeton University Press, Manchester, Princeton, 1989.