Dynamical Equivalence and Linear Conjugacy of Chemical Reaction Networks: New Results and Methods

Matthew D. Johnston[1]**, David Siegel**[1] **and Gábor Szederkényi**[2,3]

[1] *Department of Applied Mathematics, University of Waterloo,*
Waterloo, Ontario, Canada N2L 3G1

[2] *(Bio)Process Engineering Group, IIM-CSIC, Spanish National Research Council,*
C/Eduardo Cabello, 6, 36208 Vigo, Spain

[3] *Computer and Automation Research Institute, Hungarian Academy of Sciences*
H-1518, P.O. Box 63, Budapest, Hungary

Abstract

In the first part of this paper, we propose new optimization-based methods for the computation of preferred (dense, sparse, reversible, detailed and complex balanced) linearly conjugate reaction network structures with mass action dynamics. The developed methods are extensions of previously published results on dynamically equivalent reaction networks and are based on mixed-integer linear programming. As related theoretical contributions we show that (i) dense linearly conjugate networks define a unique super-structure for any positive diagonal state transformation if the set of chemical complexes is given, and (ii) the existence of linearly conjugate detailed balanced and complex balanced networks do not depend on the selection of equilibrium points. In the second part of the paper it is shown that determining dynamically equivalent realizations to a network that is structurally fixed but parametrically not can also be written and solved as a mixed-integer linear programming problem. Several examples illustrate the presented computation methods.

1 Introduction

The mathematical study of chemical reaction networks is a rapidly growing field that has been applied recently to research problems in industrial chemistry, systems biology,

gene regulation, and general nonlinear systems theory, among others [10, 23, 33]. There has also been significant theoretical work in the literature on such topics as persistence [1–3, 26, 32], multistability [8, 9, 31], monotonicity [5, 6], the global attractor conjecture for complex balanced systems [1, 2, 7, 12], lumping [15, 28, 39, 40], and conjugacy of reaction networks [11, 25].

One line of research which has been attracting increased attention has been that of determining when two reaction networks exhibit the same qualitative dynamics despite disparate network structure. In [11] and [34], the authors complete the question of what network structures can give rise to the same set of governing differential equations and therefore exhibit identical dynamics. This work was extended in [25] to networks which do not necessarily have the same set of differential equations but rather have trajectories related by a non-trivial linear transformation. Similar ground has also been touched in the papers [28] and [15] which deal with properties of linear lumpings, linear mappings which potentially reduce the dimension of the species set.

In this paper we look at the problem of algorithmically determining when a network is linearly conjugate to another network satisfying specified conditions. This problem was first addressed in [35] where the author presents a mixed-integer linear programming (MILP) algorithm capable of determining sparse and dense realizations, i.e. networks with the fewest and greatest number of reactions capable of generating the given dynamics. The algorithm was extended to complex and detailed balanced networks in [36], reversible networks in [37], weakly reversible networks in [38], and linear conjugate networks in [27], where a computationally more efficient procedure for determining weak reversibility was also presented.

We will continue in this paper the development and application of this MILP framework to linearly conjugate networks. In particular, we will give conditions which can be used to find sparse and dense linearly conjugate networks which are detailed and complex balanced, fully reversible, and which contain the greatest and fewest number of complexes. We will also expand the original MILP algorithm for finding sparse and dense realizations to find alternative realizations to a given reaction network when the network structure is fixed but the parameter values are not. Since this algorithm works without having to have the rate constants specified beforehand, this will allow us to answer questions about the reaction mechanism itself.

2 Background

In this section we present the terminology and notation relevant to chemical reaction networks and the main results from the literature upon which we will be building.

2.1 Chemical Reaction Networks

We will consider the chemical *species* or *reactants* of a network to be given by the set $\mathcal{S} = \{X_1, X_2, \ldots, X_n\}$. The combined elements on the left-hand and right-hand side of a reaction are given by linear combinations of these species. These combined terms are called *complexes* and will be denoted by the set $\mathcal{C} = \{C_1, C_2, \ldots, C_m\}$ where

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

and the α_{ij} are nonnegative integers called the *stoichiometric coefficients*. We define the reaction set to be $\mathcal{R} = \{(C_i, C_j) \mid C_i \text{ reacts to form } C_j\}$ where the property $(C_i, C_j) \in \mathcal{R}$ will more commonly be denoted $C_i \to C_j$. To each $(C_i, C_j) \in \mathcal{R}$ we will associate a positive *rate constant* $k(i, j) > 0$ and to each $(C_i, C_j) \notin \mathcal{R}$ we will set $k(i, j) = 0$. The triplet $\mathcal{N} = (\mathcal{S}, \mathcal{C}, \mathcal{R})$ will be called the *chemical reaction network*.

The above formulation naturally gives rise to a directed graph $G(V, E)$ where the set of vertices is given by $V = \mathcal{C}$, the set of directed edges is given by $E = \mathcal{R}$, and the rate constants $k(i, j)$ correspond to the weights of the edges from C_i to C_j. In the literature this has been termed the *reaction graph* of the network [24]. Since complexes may be involved in more than one reaction, as a product or a reactant, there is further graph theory we may consider. A *linkage class* is a maximally connected set of complexes, that is to say, two complexes are in the same linkage class if and only if there is a sequence of reactions in the reaction graph (of either direction) which connects them. A reaction network is called *reversible* if $C_i \to C_j$ for any $C_i, C_j \in \mathcal{C}$ implies $C_j \to C_i$. A reaction network is called *weakly reversible* if $C_i \to C_j$ for any $C_i, C_j \in \mathcal{C}$ implies there is ! some sequence of complexes such that $C_i = C_{\mu(1)} \to C_{\mu(2)} \to \cdots \to C_{\mu(l-1)} \to C_{\mu(l)} = C_j$.

A directed graph is called *strongly connected* if there exists a directed path from each vertex to every other vertex. A *strongly connected component* of a directed graph is a maximal set of vertices for which paths exists from each vertex in the set to every other vertex in the set. For a weakly reversible network, the linkage classes clearly correspond to the strongly connected components of the reaction graph.

Assuming mass-action kinetics, the dynamics of the specie concentrations over time is governed by the set of differential equations

$$\frac{d\mathbf{x}}{dt} = Y \cdot A_k \cdot \Psi(\mathbf{x}) \tag{1}$$

where $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]^T$ is the vector of reactant concentrations. The *stoichiometric matrix* Y contains entries $[Y]_{ij} = \alpha_{ji}$ and the *Kirchhoff* or *kinetics* matrix A_k is given by

$$[A_k]_{ij} = \begin{cases} -\sum_{l=1, l \neq i}^{m} k(i,l), & \text{if} \quad i = j \\ k(j,i) & \text{if} \quad i \neq j. \end{cases} \tag{2}$$

When we speak of the *structure* of a kinetics matrix, we will be referring to the distribution of positive and zero entries, which determines the network structure of the corresponding reaction graph. Finally, the *mass-action vector* $\Psi(\mathbf{x})$ is given by

$$\Psi_j(\mathbf{x}) = \prod_{i=1}^{n} x_i^{[Y]_{ij}}, \quad j = 1, \ldots, m. \tag{3}$$

2.2 Linearly Conjugate Networks

Under the assumption of mass-action kinetics, it is possible for the trajectories of two reaction networks $\mathcal{N}$ and $\mathcal{N}'$ to be related by a linear transformation and therefore share many of the same qualitative features (e.g. number and stability of equilibria, persistence/extinction of species, dimensions of invariant spaces, etc.). This phenomenon was termed *linear conjugacy* in [25].

For completeness, we include the formal definition of linear conjugacy as presented in [25]. We will let $\Phi(\mathbf{x}_0, t)$ denote the flow of (1) associated with $\mathcal{N}$ and $\Psi(\mathbf{x}_0, t)$ denote the flow of (1) associated with $\mathcal{N}'$.

Definition 2.1. *We will say two chemical reaction networks $\mathcal{N}$ and $\mathcal{N}'$ are **linearly conjugate** if there exists a linear function $\mathbf{h} : \mathbb{R}_{>0}^n \mapsto \mathbb{R}_{>0}^n$ such that $\mathbf{h}(\Phi(\mathbf{x}_0, t)) = \Psi(\mathbf{h}(\mathbf{x}_0), t)$ for all $\mathbf{x}_0 \in \mathbb{R}_{>0}^n$.*

It is known that linear transformations $\mathbf{h} : \mathbb{R}_{>0}^n \mapsto \mathbb{R}_{>0}^n$ can consist of at most positive scaling and reindexing of coordinates (Lemma 3.1, [25]). Linear conjugacy has been subsequently studied from a computational point of view in [27].

Linear conjugacy is a generalization of the concept of *dynamical equivalence* whereby two reaction networks with different topological network structure can generate the same exact set of differential equations (1). Two dynamically equivalent networks $\mathcal{N}$ and $\mathcal{N}'$

are said to be alternate *realizations* of the kinetics (1), although it is sometimes preferable to say that $\mathcal{N}$ is an alternative realization of $\mathcal{N}'$ or vice-versa. Since the case of two networks being realizations of the same kinetics is encompassed as a special case of linear conjugacy taking the transformation to be the identity, we will focus on linearly conjugate networks.

In general practice, we are given a network $\mathcal{N}$ and asked to determine its dynamical behaviour. This is often a challenging problem; however, we may notice that the network $\mathcal{N}$ *behaves like* one from a well-studied class of networks and therefore suspect a relationship which preserves key qualitative aspects of the dynamics. The theory of linear conjugacy can provide a powerful tool in analyzing such networks. If the network can be shown to be linearly conjugate to a network $\mathcal{N}'$ from the class of networks with understood dynamics, the dynamics of $\mathcal{N}'$ are transferred to $\mathcal{N}$.

This raises the question of how to find a linearly conjugate network $\mathcal{N}'$ when only the original network $\mathcal{N}$ is given. This was studied in [27] where the authors built upon the linear programming algorithm introduced in [35]. We can impose that a network $\mathcal{N}'$ be linearly conjugate to our given network $\mathcal{N}$ with the set of linear constraints

$$
\textbf{(LC)} \quad \left\{
\begin{array}{l}
Y \cdot A_b = T^{-1} \cdot M \\
\displaystyle\sum_{i=1}^{m} [A_b]_{ij} = 0, \quad j = 1, \ldots, m \\
[A_b]_{ij} \geq 0, \quad i, j = 1, \ldots, m, \ i \neq j \\
[A_b]_{ii} \leq 0, \quad i = 1, \ldots, m \\
\epsilon \leq c_j \leq 1/\epsilon, \quad j = 1, \ldots, n
\end{array}
\right.
\tag{4}
$$

where $0 < \epsilon \ll 1$, and the matrices $M \in \mathbb{R}^{n \times m}$ and $T \in \mathbb{R}^{n \times n}$ are given by:

$$
M = Y \cdot A_k, \text{ and} \tag{5}
$$

$$
T = \operatorname{diag}\{\mathbf{c}\}. \tag{6}
$$

The kinetics matrix for the network $\mathcal{N}'$ can by constructed from $A_b \in \mathbb{R}^{m \times m}$ and $\mathbf{c} \in \mathbb{R}^{n}_{>0}$ by the relation

$$
A'_k = A_b \cdot \operatorname{diag}\{\Psi(\mathbf{c})\}. \tag{7}
$$

Finding a network satisfying (4) and then solving (7) is sufficient to determine a network $\mathcal{N}'$ which is linearly conjugate to $\mathcal{N}$ via the transformation $\mathbf{h}(\mathbf{x}) = T^{-1}\mathbf{x}$.

2.3 Sparse and Dense Linearly Conjugate Networks

In order to place the problem within a linear programming framework, we need to choose an objective function to optimize. An appropriate choice of such a function is not obvious and may vary depending on the intended application.

One particularly intuitive choice, which was introduced in [35] and has been widely used since, is to search for networks $\mathcal{N}'$ with the fewest and greatest number of reactions (*sparse* and *dense* networks, respectively). A sparse (respectively, dense) linearly conjugate network is given by a matrix A'_k satisfying (4) with the most (respectively, least) off-diagonal entries which are zeroes. Since the structure of A'_k and A_b are the same, a correspondence between the non-zero off-diagonal entries in A'_k and a positive integer value can be made by considering the binary variables $\delta_{ij} \in \{0, 1\}$ which will keep track of whether a reaction is 'on' or 'off', i.e. we have

$$\delta_{ij} = 1 \iff [A_b]_{ij} > \epsilon, \quad i, j = 1, \ldots, m, \quad i \neq j$$

where $0 < \epsilon \ll 1$ is sufficiently small and can be chosen the same as in (4), and where the symbol '$\iff$' denotes the logical relation 'if and only if'. These proposition logic constraints for the structure of a network can then be formulated as the following linear mixed-integer constraints (see, for example, [30]):

$$\textbf{(S)} \quad \left\{ \begin{array}{l} 0 \leq [A_b]_{ij} - \epsilon \delta_{ij}, \quad i, j = 1, \ldots, m, \quad i \neq j \\ 0 \leq -[A_b]_{ij} + u_{ij}\delta_{ij}, \quad i, j = 1, \ldots, m, \quad i \neq j \\ \delta_{ij} \in \{0, 1\}, \quad i, j = 1, \ldots, m, i \neq j, \end{array} \right. \quad (8)$$

where $u_{ij} > 0$ for $i, j = 1, \ldots, m, i \neq j$, are appropriate upper bounds for the reaction rate coefficients. The number of reactions present in the network corresponding to A_k is then given by the sum of the δ_{ij}'s so that the problem of determining a sparse network corresponds to satisfying the objective function

$$\textbf{(Sparse)} \quad \left\{ \quad \text{minimize} \quad \sum_{i,j=1, i \neq j}^{m} \delta_{ij} \right. \quad (9)$$

over the constraint sets (4) and (8). Finding a dense network corresponds to maximizing the same function, which can also be stated as a minimization problem as

$$\textbf{(Dense)} \quad \left\{ \quad \text{minimize} \quad \sum_{i,j=1, i \neq j}^{m} -\delta_{ij}. \right. \quad (10)$$

It is known that, for trivial linear conjugacies, the structure of the dense realization contains the structure of all other trivial linear conjugacies as a subset (Theorem 3.1, [37]). We now prove the comparable result for non-trivial linear conjugacies.

Theorem 2.1. *Let $\mathcal{N}$ be a chemical reaction network. Suppose that the reaction network $\mathcal{N}'$ is linearly conjugate to $\mathcal{N}$ and dense. Suppose that $\mathcal{N}''$ is also linearly conjugate to $\mathcal{N}$. Then the directed unweighted graph of $\mathcal{N}''$ is a subset of the directed unweighted graph of $\mathcal{N}'$.*

Proof. Assume $\mathcal{N}'$ and $\mathcal{N}''$ are linearly conjugate to $\mathcal{N}$, $\mathcal{N}'$ is dense in the space of networks which are linearly conjugate to $\mathcal{N}$, and $\mathcal{N}''$ contains a reaction not contained in $\mathcal{N}'$.

Since both $\mathcal{N}'$ and $\mathcal{N}''$ are linearly conjugate to $\mathcal{N}$, we have by (4) that

$$Y \cdot A_k = T' \cdot Y \cdot A_b'$$

and

$$Y \cdot A_k = T'' \cdot Y \cdot A_b''$$

where $T' = \operatorname{diag}\{\mathbf{c}'\}$, $\mathbf{c}' \in \mathbb{R}_{>0}^n$ and A_b' correspond to $\mathcal{N}'$, $\mathbf{c}'' \in \mathbb{R}_{>0}^n$ and $T'' = \operatorname{diag}\{\mathbf{c}''\}$ and A_b'' correspond to $\mathcal{N}''$.

Now consider $T' \cdot Y \cdot A_b''$. We have

$$T' \cdot Y \cdot A_b'' = T' \cdot (T'')^{-1} \cdot T'' \cdot Y \cdot A_b'' = Q \cdot Y \cdot A_k$$

where $Q = T' \cdot (T'')^{-1}$. Consequently, we have

$$T' \cdot Y \cdot A_b'' + T' \cdot Y \cdot A_b' = (Q + I) \cdot Y \cdot A_k. \tag{11}$$

On the other hand, we have

$$T' \cdot Y \cdot A_b'' + T' \cdot Y \cdot A_b' = T' \cdot Y \cdot (A_b'' + A_b') = T' \cdot Y \cdot A_b''' \tag{12}$$

where $A_b''' = A_b' + A_b''$. Combining (11) and (12) gives

$$Y \cdot A_k = (Q + I)^{-1} \cdot T' \cdot Y \cdot A_b''' = T''' \cdot Y \cdot A_b'''$$

where $T''' = (T' \cdot (T'')^{-1} + I)^{-1} \cdot T'$.

Since T', T'' and I are diagonal matrices with positive entries on the diagonal, so is T'''. This means that the network $\mathcal{N}'''$ corresponding to $A_b''' = A_b' + A_b''$ is linearly conjugate to $\mathcal{N}$. We can readily see, however, that $\mathcal{N}'''$ contains all of the reactions in both $\mathcal{N}'$ and $\mathcal{N}''$. If $\mathcal{N}''$ contains a reaction not contained in $\mathcal{N}'$ then $\mathcal{N}'''$ clearly has more reactions than $\mathcal{N}'$ which contradicts the assumption that $\mathcal{N}'$ is dense in the space of networks which are linearly conjugate to $\mathcal{N}$. The result follows. $\qquad\square$

The following result follows immediately.

Corollary 2.1. *Let $\mathcal{N}$ be a chemical reaction network. Then the structure of the unweighted directly graph of the dense reaction network $\mathcal{N}'$ which is linearly conjugate to $\mathcal{N}$ is unique.*

Proof. This follows directly from Theorem 2.1. $\square$

3 Computational Extensions of Linearly Conjugate Networks

In this section we will extend the optimization framework introduced in Section 2.3 to include complex balanced, reversible and detailed balanced networks, and to search for networks with the greatest and fewest number of complexes.

3.1 Weakly Reversible Networks

Weakly reversible networks are a particular important class of reaction networks because strong properties are known about their dynamics and equilibrium concentrations. Under a supplemental condition on the rate constants, weakly reversible networks are known to have complex balanced equilibrium concentrations and therefore exhibit all of the dynamical properties normally reserved for these networks [17, 22] (see Section 3.3 for further discussion of complex balanced networks).

Consequently, they are a primary candidate for the type of network we would like to find. The problem of determining if and when a chemical reaction network $\mathcal{N}$ is linearly conjugate to a weakly reversible network $\mathcal{N}'$ was first considered in [38] and further refined in [27]. For convenience, we briefly recall the constraints published in [27] that guarantee the weak reversibility of the linearly conjugate network $\mathcal{N}'$:

$$\textbf{(WR)} \quad \left\{ \begin{array}{l} \displaystyle\sum_{i=1,i\neq j}^{m} [\tilde{A}_k]_{ij} = \sum_{i=1,i\neq j}^{m} [\tilde{A}_k]_{ji}, \quad j=1,\ldots,m \\ [\tilde{A}_k]_{ij} \geq 0, \quad i,j=1,\ldots,m, \; i\neq j \end{array} \right. \tag{13}$$

where $\tilde{A}_k$ is an auxiliary Kirchhoff matrix with the same structure as A'_k with appropriately scaled columns such that its kernel contains the m-dimensional vector of all ones. In order to guarantee that the matrix $\tilde{A}_k$ has the same structure as A'_k and A_b we also require that

$$\textbf{(WR-S)} \quad \left\{ \begin{array}{l} 0 \leq [\tilde{A}_k]_{ij} - \epsilon\delta_{ij}, \quad i,j=1,\ldots,m, \; i\neq j \\ 0 \leq -[\tilde{A}_k]_{ij} + u_{ij}\delta_{ij}, \quad i,j=1,\ldots,m, \; i\neq j. \end{array} \right. \tag{14}$$

3.2 Reversible Networks

In [36], an algorithm was presented which was capable of determining reversible reaction networks which are trivially linearly conjugate to a given reaction network. In this section, we present a simplified methodology and apply it to non-trivial linear conjugacies.

We recall that a network is reversible if $\mathcal{C}_i \to \mathcal{C}_j$ for any $\mathcal{C}_i, \mathcal{C}_j \in \mathcal{C}$ implies $\mathcal{C}_j \to \mathcal{C}_i$. For the network $\mathcal{N}'$, this is equivalent to the condition

$$[A'_k]_{ij} > \epsilon \iff [A'_k]_{ji} > \epsilon$$

for some sufficient small $0 < \epsilon \ll 1$. This is in turn equivalent to

$$\delta_{ij} = 1 \iff \delta_{ji} = 1$$

where $\delta_{ij} \in \{0, 1\}$, $i, j = 1, \ldots, m$, $i \neq j$, as in Section 2.3. It follows that we can restrict our search space to reversible networks with the constraint set

$$\textbf{(Rev)} \quad \left\{ \begin{array}{l} \delta_{ij} - \delta_{ji} = 0, \\ i, j = 1, \ldots, m, \ \ i < j. \end{array} \right. \tag{15}$$

A sparse or dense reversible network which is linearly conjugate to $\mathcal{N}$ can be found by optimizing (9) or (10), respectively, over the constraint sets (4), (8) and (15).

3.3 Complex Balanced Systems

A particularly important class of chemical reaction networks are the complex balanced networks introduced in [24].

Definition 3.1. *An equilibrium concentration* $\mathbf{x}^* \in \mathbb{R}^n_{>0}$ *of the chemical reaction network* $\mathcal{N}$ *is a **complex balanced equilibrium concentration** if*

$$A_k \cdot \Psi(\mathbf{x}^*) = \mathbf{0}. \tag{16}$$

The network $\mathcal{N}$ *is called **complex balanced** if every equilibrium concentration* $\mathbf{x}^* \in \mathbb{R}^n_{>0}$ *is a complex balanced equilibrium concentration.*

Many strong properties are known about complex balanced networks. In particular, it is known that complex balanced networks permit exactly one positive equilibrium concentration in each invariant space of the network and that this equilibrium concentration is locally asymptotically stable (Lemma 4C and Theorem 6A, [24]). Complex balanced

systems are also known to be weakly reversible so that they are a subset of the weakly reversible networks considered in Section 3.1 (Theorem 2B, [22]).

The following result shows complex balancing is a system property depending on the structure and parameters of the network and not on the chosen equilibrium concentration.

Theorem 3.1 (Theorem 6A, [24]). *If a chemical reaction network $\mathcal{N}$ is complex balanced at an equilibrium concentration $\mathbf{x}^* \in \mathbb{R}_{>0}^n$ then it is complex balanced at all of its equilibrium concentrations.*

It should be noted that the complex balancing of a network is still dependent on the choice of rate constants. It is possible for a reaction network to be complex balanced for some choices of rate constants and not for others.

In [36], an algorithm was presented which was capable of determining sparse and dense complex balanced networks which are trivially linearly conjugate to a given network $\mathcal{N}$. This method required determining an equilibrium value $\mathbf{x}^* \in \mathbb{R}_{>0}^n$ of the network $\mathcal{N}$ and then imposing the condition

$$A_k' \cdot \Psi(\mathbf{x}^*) = \mathbf{0}$$

on $\mathcal{N}'$ in accordance with (16). In this section, we extend these results to include non-trivial linearly conjugate networks.

Suppose that $\mathcal{N}$ and $\mathcal{N}'$ are linearly conjugate via the transformation $\mathbf{y} = T^{-1}\mathbf{x}$. In order to guarantee the network $\mathcal{N}'$ is complex balanced, according to (16) we require that

$$A_k' \cdot \Psi(\mathbf{y}^*) = \mathbf{0}. \tag{17}$$

Since the equilibrium concentrations of $\mathcal{N}$ and $\mathcal{N}'$ are related by the transformation $\mathbf{y}^* = T^{-1}\mathbf{x}^*$, we have that the left-hand side of (17) can be rewritten

$$A_k' \cdot \Psi(\mathbf{y}^*) = A_k' \cdot \Psi(T^{-1}\mathbf{x}^*) = A_k' \cdot \text{diag}\left\{\Psi(\mathbf{c})\right\}^{-1} \cdot \Psi(\mathbf{x}^*) = A_b \cdot \Psi(\mathbf{x}^*)$$

where we have made use of the form of the kinetics matrix of $\mathcal{N}'$ according to (7). The condition for complex balancing of the linearly conjugate network $\mathcal{N}'$ is therefore

$$\textbf{(CB)} \quad \begin{cases} A_b \cdot \Psi(\mathbf{x}^*) = \mathbf{0} \\ M \cdot \Psi(\mathbf{x}^*) = \mathbf{0} \\ \mathbf{x}^* \in \mathbb{R}_{>0}^n \end{cases} \tag{18}$$

where M is as in (5). A sparse or dense complex balanced network which is linearly conjugate to $\mathcal{N}$ can be found by optimizing (9) or (10), respectively, over the constraint sets (4), (8) and (18).

It should be noted that the optimization algorithm is less computationally exhausting than the corresponding algorithm for weak reversibility (13). This is because the matrix $\tilde{A}_k$ required in the general weakly reversible case is not required in the complex balancing condition; rather, it is sufficient to use the matrix A_b. Consequently, there are fewer decision variables in the complex balancing algorithm. The pre-step that a $\mathbf{x}^* \in \mathbb{R}_{>0}^n$ be found satisfying $Y \cdot A_k \cdot \Psi(\mathbf{x}^*) = \mathbf{0}$ may off-set this advantage, however, depending on the difficulty in solving $Y \cdot A_k \cdot \Psi(\mathbf{x}^*) = \mathbf{0}$.

It is unclear how the outcome of the algorithm depends on the choice of equilibrium concentration $\mathbf{x}^* \in \mathbb{R}_{>0}^n$, which in general is not unique. The following result clarifies this dependence.

Theorem 3.2. *Suppose $\mathcal{N}$ is linearly conjugate to $\mathcal{N}'$ with transformation matrix $T = \mathrm{diag}\{\mathbf{c}\}$ where $\mathbf{c} \in \mathbb{R}_{>0}^n$ and suppose $\mathcal{N}'$ is complex balanced at $\mathbf{y}^* = T^{-1}\mathbf{x}^*$ where $\mathbf{x}^* \in \mathbb{R}_{>0}^n$ and $Y \cdot A_k \cdot \Psi(\mathbf{x}^*) = \mathbf{0}$. Then $\mathcal{N}'$ is complex balanced at $\bar{\mathbf{y}}^* = T^{-1}\bar{\mathbf{x}}^*$ for all $\bar{\mathbf{x}}^* \in \mathbb{R}_{>0}^n$ satisfying $Y \cdot A_k \cdot \Psi(\bar{\mathbf{x}}^*) = \mathbf{0}$.*

Proof. Suppose trajectories of $\mathcal{N}'$ are related to trajectories of $\mathcal{N}$ by the relationship $\mathbf{y} = T^{-1}\mathbf{x}$ where $T = \mathrm{diag}\{\mathbf{c}\}$ and $\mathbf{c} \in \mathbb{R}_{>0}^n$.

Suppose that $\mathcal{N}'$ is complex balanced at $\mathbf{y}^* = T^{-1}\mathbf{x}^*$ where $\mathbf{x}^* \in \mathbb{R}_{>0}^n$ is an equilibrium concentration of $\mathcal{N}$. It follows from Theorem 3.1 that

$$Y \cdot A'_k \cdot \Psi(\mathbf{y}) = \mathbf{0} \quad \Longrightarrow \quad A'_k \cdot \Psi(\mathbf{y}) = \mathbf{0}. \tag{19}$$

Now consider an arbitrary equilibrium concentration $\bar{\mathbf{x}}^* \in \mathbb{R}_{>0}^n$ of $\mathcal{N}$. Since $\mathcal{N}$ and $\mathcal{N}'$ are linearly conjugate, it follows that $Y \cdot A_k = T \cdot Y \cdot A'_k \cdot \mathrm{diag}\{\Psi(\mathbf{c})\}^{-1}$. It follows that we have

$$\mathbf{0} = Y \cdot A_k \cdot \Psi(\bar{\mathbf{x}}^*) = T \cdot Y \cdot A'_k \cdot \mathrm{diag}\{\Psi(\mathbf{c})\}^{-1} \cdot \Psi(T\bar{\mathbf{y}}^*)$$

$$= T \cdot Y \cdot A'_k \cdot \Psi(\bar{\mathbf{y}}^*).$$

It follows by the structure of T that $Y \cdot A'_k \cdot \Psi(\bar{\mathbf{y}}^*) = \mathbf{0}$. From (19) we have that $A'_k \cdot \Psi(\bar{\mathbf{y}}^*) = \mathbf{0}$. In other words, $\mathcal{N}'$ is complex balanced at $\bar{\mathbf{y}}^*$ and we are done.

$\square$

This result shows that when imposing the complex balancing constraint (18) on $\mathcal{N}'$, it does not matter which equilibrium concentration of $\mathcal{N}$ we choose. The feasible set of solutions (i.e. admissible networks) is the same.

3.4 Detailed Balanced Systems

In [36], the authors present an algorithm for determining detailed balanced networks which are trivially linearly conjugate to a given network. In this section we extend this algorithm to non-trivial linear conjugacies.

Definition 3.2. *An equilibrium concentration* $\mathbf{x}^* \in \mathbb{R}^n_{>0}$ *of the chemical reaction network* $\mathcal{N}$ *is a* **detailed balanced** *equilibrium concentration if*

$$[A_k]_{ij}\Psi_j(\mathbf{x}^*) = [A_k]_{ji}\Psi_i(\mathbf{x}^*), \quad \forall\, i,j = 1, \ldots, m, \;\; i \neq j. \tag{20}$$

The network $\mathcal{N}$ *is called* **detailed balanced** *if every equilibrium concentration* $\mathbf{x}^* \in \mathbb{R}^n_{>0}$ *is a detailed balanced equilibrium concentration.*

In other words, an equilibrium concentration $\mathbf{x}^* \in \mathbb{R}^n_{>0}$ is detailed balanced if the flow across each reaction is balanced by the flow across an opposite reaction at $\mathbf{x}^*$.

Suppose that $\mathcal{N}$ and $\mathcal{N}'$ are linearly conjugate via the transformation $\mathbf{y} = T^{-1}\mathbf{x}$. In order to guarantee the network $\mathcal{N}'$ is detailed balanced, according to (20) we require that

$$\mathrm{diag}\left\{\Psi(\mathbf{y}^*)\right\} \cdot (A'_k)^T = A'_k \cdot \mathrm{diag}\left\{\Psi(\mathbf{y}^*)\right\}. \tag{21}$$

Since the equilibrium concentrations of $\mathcal{N}$ and $\mathcal{N}'$ are related by the transformation $\mathbf{y}^* = T^{-1}\mathbf{x}^*$, we have that

$$A'_k \cdot \mathrm{diag}\left\{\Psi(\mathbf{y}^*)\right\} = A'_k \cdot \mathrm{diag}\left\{\Psi(\mathbf{c})\right\}^{-1} \cdot \mathrm{diag}\left\{\Psi(\mathbf{x}^*)\right\}$$
$$= A_b \cdot \mathrm{diag}\left\{\Psi(\mathbf{x}^*)\right\}$$

where we have made use of the form of the kinetics matrix of $\mathcal{N}'$ according to (7). The condition for detailed balancing of the linearly conjugate network $\mathcal{N}'$ is therefore

$$\textbf{(DB)} \quad \left\{ \begin{array}{l} \mathrm{diag}\left\{\Psi(\mathbf{x}^*)\right\} \cdot A_b^T = A_b \cdot \mathrm{diag}\left\{\Psi(\mathbf{x}^*)\right\} \\ M \cdot \Psi(\mathbf{x}^*) = \mathbf{0} \\ \mathbf{x}^* \in \mathbb{R}^n_{>0} \end{array} \right. \tag{22}$$

where M is as in (5). A sparse or dense detailed balanced network which is linearly conjugate to $\mathcal{N}$ can be found by optimizing (9) or (10), respectively, over the constraint sets (4), (8) and (22). Since the analogous result to Theorem 3.2 holds as a consequence of detailed balanced systems being a subset of complex balanced networks, we do not prove it here.

3.5   Minimal and Maximal Number of Complexes

We can also adapt to non-trivial linear conjugacies the algorithm introduced in [36] for determining a network with the fewest or greatest number of complexes from a fixed complex set which is trivially linearly conjugate to a given network $\mathcal{N}$.

In order to count the number of complexes in the network, we introduce the binary variables $\delta_i \in \{0,1\}$, $i, j = 1, \ldots, m$, and consider the logical equations

$$\delta_i = 1 \iff \sum_{\substack{j_1=1 \\ j_1 \neq i}}^{m} [A_k]_{ij_1} + \sum_{\substack{j_2=1 \\ j_2 \neq i}}^{m} [A_k]_{j_2 i} > 0 \tag{23}$$

for $i = 1, \ldots, m$. In other words, δ_i takes on the value of one if and only if there is a reaction to or from the complex $\mathcal{C}_i$ in the network; otherwise, it takes the value zero. For computational purposes, we reconsider (23) as

$$\delta_i = 1 \iff \sum_{\substack{j_1=1 \\ j_1 \neq i}}^{m} [A_k]_{ij_1} + \sum_{\substack{j_2=1 \\ j_2 \neq i}}^{m} [A_k]_{j_2 i} \geq \epsilon \tag{24}$$

where $0 < \epsilon \ll 1$. The linear constraints required to substantiate (24) are

$$(\textbf{Comp}) \begin{cases} 0 \leq \sum_{\substack{j_1=1 \\ j_1 \neq i}}^{m} [A_k]_{ij_1} + \sum_{\substack{j_2=1 \\ j_2 \neq i}}^{m} [A_k]_{j_2 i} - \delta_i \epsilon \\ 0 \leq - \sum_{\substack{j_1=1 \\ j_1 \neq i}}^{m} [A_k]_{ij_1} - \sum_{\substack{j_2=1 \\ j_2 \neq i}}^{m} [A_k]_{j_2 i} \\ \qquad\qquad + \left(\sum_{\substack{j_1=1 \\ j_1 \neq i}}^{m} u_{ij_1} + \sum_{\substack{j_2=1 \\ j_2 \neq i}}^{m} u_{j_2 i} \right) \delta_i \\ \delta_i \in \{0,1\}, i = 1, \ldots, m. \end{cases} \tag{25}$$

We can now determine a network with the fewest or greatest number of complexes by optimizing the functions

$$(\textbf{Min}) \left\{ \quad \text{minimize} \quad \sum_{i=1}^{m} \delta_i \right. \tag{26}$$

or

$$(\textbf{Max}) \left\{ \quad \text{minimize} \quad - \sum_{i=1}^{m} \delta_i \right. \tag{27}$$

respectively over the constraint sets (4) and (25). Further constraints can be imposed to restrict ourselves to specific classes of systems (e.g. complex balanced systems, reversible networks, etc.) although care has to be taken to ensure the structural constraints are still satisfied.

3.6 Examples

In this section, we present a few examples which illustrate the methodologies outlined so far.

Example 1: Consider the kinetic scheme

$$\dot{x}_1 = x_1 x_2^2 - 2x_1^2 + x_1 x_3^2$$
$$\dot{x}_2 = -x_1^2 x_2^2 + x_1 x_3^2 \tag{28}$$
$$\dot{x}_3 = x_1^2 - 3x_1 x_3^2$$

introduced in [27]. In that paper, it was shown that the kinetics could be generated by a reaction network involving the complex set

$$C_1 = X_1 + 2X_2, C_2 = 2X_1 + 2X_2, C_3 = 2X_1 + X_2,$$

$$C_4 = 2X_1, C_5 = X_1, C_6 = 2X_1 + X_3, C_7 = X_1 + 2X_3$$

$$C_8 = 2X_1 + 2X_3, C_9 = X_1 + X_2 + 2X_3, C_{10} = X_1 + X_3$$

and that (28) has dynamics which is linearly conjugate to those generated by the sparse weakly reversible network given in Figure 1(a) (conjugacy constants $c_1 = 20, c_2 = 2, c_3 = 5$) and the dense weakly reversible network given in Figure 1(b) (conjugacy constants $c_1 = 20/3, c_2 = 20/33, c_3 = 5/3$).
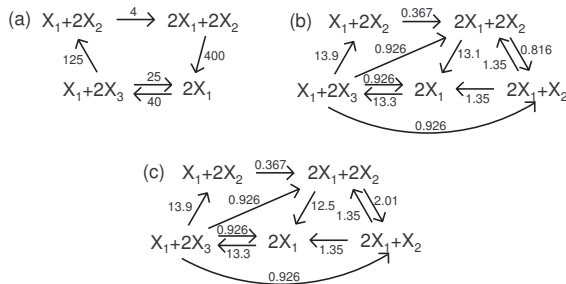

Figure 1: Weakly reversible networks which are linearly conjugate to a network with the kinetics (28). The network in (a) is sparse while the networks in (b) and (c) are dense. The networks (a) and (c) are also complex balanced. (The parameter values in (b) and (c) have been rounded to three significant figures.)

The network in Figure 1(a) is complex balanced as a consequence of it being a zero deficiency weakly reversible network. It can be easily checked, however, that the network

in Figure 1(b) is not complex balanced. We might wonder, therefore, what running the algorithm for a dense complex balanced network which is linearly conjugate to a network generating the kinetics (28) would produce.

Numerically, we can determine that an equilibrium concentration of (28) is $(x_1^*, x_2^*, x_3^*) = (0.2, 0.577350269, 0.258198889)$. Running GLPK for a sparse network (9) over the constraints (4), (8) and (18) gives the network given in Figure 1(c) (conjugacy constants $c_1 = 20/3, c_2 = 20/33, c_3 = 5/3$). We notice that this network has the same structure as the weakly reversible network in Figure 1(b) and, furthermore, only differs in two rate constant values. It can be checked, however, that (c) is complex balanced while (b) is not.
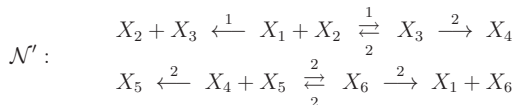
Example 2: Consider the kinetic scheme

$$\begin{array}{ll} \dot{x}_1 = -2x_1x_2 + 2x_3 + 2x_6 & \dot{x}_4 = x_3 - x_4x_5 + x_6 \\ \dot{x}_2 = -x_1x_2 + 2x_3 & \dot{x}_5 = -2x_4x_5 + 4x_6 \\ \dot{x}_3 = 2x_1x_2 - 4x_3 & \dot{x}_6 = x_4x_5 - 2x_6. \end{array} \tag{29}$$

Using the indexing scheme introduced in [20] and more recently applied in the papers [36] and [27] we can construct a chemical reaction network capable of generating the dynamics (29) under the assumption of mass-action kinetics (1) which involves the complexes

$$C_1 = X_1 + X_2, C_2 = X_2, C_3 = X_3, C_4 = X_1 + X_3,$$

$$C_5 = X_6, C_6 = X_1 + X_6, C_7 = X_1, C_8 = X_2 + X_3,$$

$$C_9 = X_1 + X_2 + X_3, C_{10} = \emptyset, C_{11} = X_3 + X_4$$

$$C_{12} = X_4 + X_5, C_{13} = X_5, C_{14} = X_6, C_{15} = X_4 + X_6,$$

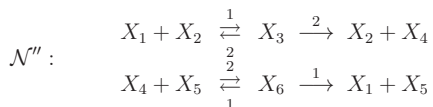$$C_{16} = X_4, C_{17} = X_5 + X_6, C_{18} = X_4 + X_5 + X_6.$$

It seems less than desirable, however, to consider a network of 18 complexes given the simplistic dynamics (29). We might wonder if there is a network with fewer complexes. Optimizing (26) over the constraint sets (4) and (25) gives the network

$$\mathcal{N}' : \quad \begin{array}{c} X_2 + X_3 \xleftarrow{1} X_1 + X_2 \underset{2}{\overset{1}{\rightleftarrows}} X_3 \xrightarrow{2} X_4 \\ \\ X_5 \xleftarrow{2} X_4 + X_5 \underset{2}{\overset{2}{\rightleftarrows}} X_6 \xrightarrow{2} X_1 + X_6 \end{array}$$

and the conjugacy constants $c_1 = 1, c_2 = 2, c_3 = 2, c_4 = 1, c_5 = 4, c_6 = 2$.

In terms of understanding the qualitative dynamics of 29), this network is not particularly insightful. We might notice, however, that the net effect of all the reaction pathways leading out from X_3 is to create an X_2 and an X_4 at the expense of depleting X_3. The

complex $X_2 + X_4$, however, has not been considered in the procedure. Similarly, the reaction pathways leading out from X_6 generate an X_1 and an X_5 but the complex $X_1 + X_5$ has not been included in the procedure. We might consider appending the procedure, therefore, to include $C_{19} = X_2 + X_4$ and $C_{20} = X_1 + X_5$. Repeating the algorithm in GLPK gives the network

$$\mathcal{N}'' : \quad \begin{aligned} X_1 + X_2 \; &\overset{1}{\underset{2}{\rightleftarrows}} \; X_3 \; \overset{2}{\longrightarrow} \; X_2 + X_4 \\ X_4 + X_5 \; &\overset{2}{\underset{1}{\rightleftarrows}} \; X_6 \; \overset{1}{\longrightarrow} \; X_1 + X_5 \end{aligned}$$

and the conjugacy constants $c_1 = 1, c_2 = 1, c_3 = 2, c_4 = 1, c_5 = 2, c_6 = 1$. This is easily identified as the enzyme network

$$\begin{aligned} S_1 + E \; &\rightleftarrows \; SE \; \longrightarrow \; S_2 + E \\ S_2 + F \; &\rightleftarrows \; PF \; \longrightarrow \; S_1 + F \end{aligned}$$

where an enzyme E facilitates the transfer of a substrate S_1 into another substract S_2 and another enzyme F facilitates the transfer back. This network was considered extensively in [3] and [4]. In particular, it was shown in [4] that for all rate constant values, the network possesses within each invariant space a unique positive equilibrium concentration which is globally asymptotically stable relative to that invariant space. It follows by the properties of linearly conjugate reaction networks that (29) inherits the same qualitative dynamics.

4 Structural Dynamical Equivalence

In this section we extend the computation procedure given in Section 3 for dynamical equivalence to the case of networks $\mathcal{N}$ which are structurally fixed but have undetermined parameters.

4.1 Dynamical Equivalence

The MILP framework outlined so far requires that the rate constants be specified for the network $\mathcal{N}$. Consequently, when we search for networks which are linearly conjugate to a given network $\mathcal{N}$, we are really asking if there are networks which are linearly conjugate *for a given choice of parameter values*.

For networks where the dynamical behaviour is heavily dependent on the chosen rate constants, however, it is possible that certain behaviours are being overlooked by poor

rate constant selection. There are networks, for instance, which are known to be linearly conjugate to weakly reversible networks or complex balanced networks for certain values of the rate constants but not for others (see Examples 2 and 3 of [25]). In such cases, if the rate constants are not carefully chosen, the algorithm may overlook these networks and we would not realize that the mechanism shares characteristics with these other networks.

Therefore, we now change the problem setup by fixing only the structure of the initial network $\mathcal{N}$ but not the parameter values. We will show that this problem class can also be casted to the framework of MILP. This would remove the above mentioned limits of using a fully specified initial network model.

The conditions for dynamical equivalence, keeping the entries of both A_k and A'_k general, are

$$
\textbf{(DE)} \quad
\left\{
\begin{array}{l}
Y \cdot A'_k = Y \cdot A_k \\
\displaystyle\sum_{i=1}^{m}[A'_k]_{ij} = \sum_{i=1}^{m}[A_k]_{ij} = 0, \quad j = 1, \ldots, m \\
[A'_k]_{ij} \geq 0, [A_k]_{ij} \geq 0, \quad i, j = 1, \ldots, m, \ i \neq j \\
[A'_k]_{ii} \leq 0, [A_k]_{ii} \leq 0, \quad i = 1, \ldots, m
\end{array}
\right.
\tag{30}
$$

Note that in (30) both the off-diagonal entries of A_k and A'_k are now decision variables.

As in Section 2.3, we want to keep track of the structure of A'_k. The conditions corresponding to (8) for the matrix A'_k are

$$
\textbf{(S2)} \quad
\left\{
\begin{array}{l}
0 \leq [A'_k]_{ij} - \epsilon\delta_{ij}, \quad i, j = 1, \ldots, m, \ i \neq j \\
0 \leq -[A'_k]_{ij} + u_{ij}\delta_{ij}, \quad i, j = 1, \ldots, m, \ i \neq j \\
\delta_{ij} \in \{0, 1\}, \quad i, j = 1, \ldots, m, i \neq j,
\end{array}
\right.
\tag{31}
$$

As before, the binary variables δ_{ij} keep track of whether a reaction is in the network $\mathcal{N}'$ or not and thus are capable of counting the number of reactions in $\mathcal{N}'$.

We also, however, want to permit A_k to have a variable rate constant values within a fixed network structure. In order to fix this network structure, we introduce the binary variables $\gamma_{ij} \in \{0, 1\}$, $i, j = 1, \ldots, m$, $i \neq j$, and the logical equations

$$
\gamma_{ij} = 1 \iff [A_k]_{ij} > \epsilon, \quad i, j = 1, \ldots, m, i \neq j
\tag{32}
$$

for some $0 < \epsilon \ll 1$. In other words, the γ_{ij}'s keep track of whether the reaction $\mathcal{C}_j \to \mathcal{C}_i$ is in the network $\mathcal{N}$. The conditions required to allow the entries of A_k to vary independently within this pre-defined structure are

$$
\textbf{(Ind)} \quad
\left\{
\begin{array}{l}
0 \leq [A_k]_{ij} - \epsilon\gamma_{ij}, \quad i, j = 1, \ldots, m, \ i \neq j \\
0 \leq -[A_k]_{ij} + u_{ij}\gamma_{ij}, \quad i, j = 1, \ldots, m, \ i \neq j
\end{array}
\right.
\tag{33}
$$

where

$$\gamma_{ij} = \begin{cases} 1, & \text{if } (\mathcal{C}_j, \mathcal{C}_i) \in \mathcal{R} \\ 0, & \text{otherwise.} \end{cases} \tag{34}$$

Further constraints can be implemented to search through subspaces of the parameter spaces for alternative realizations. For instance, if we suspect the reaction rate for the reaction $\mathcal{C}_{j_1} \to \mathcal{C}_{i_1}$ is slaved to that of $\mathcal{C}_{j_2} \to \mathcal{C}_{i_2}$ we can add $[A_k]_{i_1 j_1} = [A_k]_{i_2 j_2}$ to the procedure, etc.

The conditions (30), (31) and (33) can be combined with the structural conditions for reversibility (15) and weak reversibility (13 and 14) and the objective functions (9) and (10) to search over the parameter space of $\mathcal{N}$ for sparse and dense alternative realizations $\mathcal{N}'$ which satisfy these further structural constraints.

4.2 Complex Balanced Realizations

It is also desirable to explore the parameter space of $\mathcal{N}$ for alternative realizations $\mathcal{N}'$ which are complex balanced. The linear constraints (22) and (18), however, cannot be used in the parameter-independent case since the required equilibrium concentrations $\mathbf{x}^* \in \mathbb{R}^n_{>0}$ depend on the rate constants for $\mathcal{N}$ which are not specified.

We might expect, however, since all weakly reversible networks are complex balanced for some choice of rate constants, that if a network $\mathcal{N}$ has a weakly reversible alternative realization $\mathcal{N}'$ for some other choice of rate constants then it also has a complex balanced alternative realization $\mathcal{N}''$ for some choice of rate constants. That is to say, in the parameter-independent optimization procedure weak reversibility is sufficient to demonstrate complex balancing. The main result of this subsection guarantees this (Theorem 4.3).

First, however, we need the following results about weakly reversible networks.

Theorem 4.1 (Theorem 3.1, [19]; Proposition 4.1, [16]). *Let A_k be a kinetics matrix and let Λ_i, $i = 1, \ldots, \ell$, denote the support of the ith linkage class. Then the reaction graph corresponding to A_k is weakly reversible if and only if there is a basis of $\ker(A_k)$, $\{\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(\ell)}\}$, such that, for $i = 1, \ldots, \ell$,*

$$\mathbf{b}^{(i)} = \begin{cases} b_j^{(i)} > 0, & j \in \Lambda_i \\ b_j^{(i)} = 0, & j \notin \Lambda_i. \end{cases}$$

Theorem 4.2 (Theorem 1, [13]). *Under the assumption of mass-action kinetics, weakly reversible chemical reaction networks possess at least one positive equilibrium concentration within each positive invariant space of the system.*

An immediate consequence of Theorem 4.1 is that a network is weakly reversible if and only if there is a vector $\mathbf{b} \in \mathbb{R}_{>0}^n$ in the kernel of A_k. We will exploit this fact in the next result.

Theorem 4.3. *Suppose there is a choice of rate constants such that the network $\mathcal{N}$ is dynamically equivalent to $\mathcal{N}'$ and $\mathcal{N}'$ is weakly reversible. Then there exists a choice of rate constants such that the network $\mathcal{N}$ is dynamically equivalent to $\mathcal{N}''$ where $\mathcal{N}''$ is complex balanced. Furthermore $\mathcal{N}''$ can be selected to have the same structure as $\mathcal{N}'$.*

Proof. Let A_k be the kinetics matrix associated with $\mathcal{N}$ and A_k' be the kinetics matrix associated with $\mathcal{N}'$, and suppose that $\mathcal{N}'$ is weakly reversible. Let $\mathbf{b} \in \mathbb{R}_{>0}^n$ denote the positive vector in $\ker(A_k)$ guaranteed to exist by Theorem 4.1 and $\mathbf{x}^* \in \mathbb{R}_{>0}^n$ be any positive equilibrium concentration of (1) guaranteed to exist by Theorem 4.2.

We now define a new network $\mathcal{N}''$ with the associated kinetics matrix

$$A_k'' = A_k' \cdot \operatorname{diag}\left\{ \frac{\mathbf{b}}{\Psi(\mathbf{x}^*)} \right\} \tag{35}$$

where we define vector division to be componentwise division, i.e.

$\mathbf{x}/\mathbf{y} = [x_1/y_1, x_2/y_2, \ldots, x_n/y_n]$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Notice that all the terms in the definition of A_k'' can be determined under the assumption that $\mathcal{N}'$ is weakly reversible. We have that

$$A_k'' \cdot \Psi(\mathbf{x}^*) = A_k' \cdot \mathbf{b} = \mathbf{0}$$

since $\mathbf{b} \in \ker(A_k')$ so that $\mathcal{N}''$ is complex balanced at $\mathbf{x}^* \in \mathbb{R}_{>0}^m$. Furthermore, we have that

$$Y \cdot A_k'' = Y \cdot A_k' \cdot \operatorname{diag}\left\{ \frac{\mathbf{b}}{\Psi(\mathbf{x}^*)} \right\} = Y \cdot A_k \cdot \operatorname{diag}\left\{ \frac{\mathbf{b}}{\Psi(\mathbf{x}^*)} \right\}.$$

It is clear that $A_k \cdot \operatorname{diag}\{\mathbf{b}/\Psi(\mathbf{x}^*)\}$ has the same structure as A_k so that this corresponds to different choice of rate constants for the network $\mathcal{N}$. The result follows. $\square$

Note that, while we do not have linear constraints capable of determining complex balanced networks explicitly, we can always construct a complex balanced network from a weakly reversible or reversible network output by the optimization procedure. (Although this may not hold if further restrictions on the parameter space of $\mathcal{N}$ have been imposed.)

It is worth noting that the corresponding result for reversible and detailed balanced networks does not follow in the same manner as the proof of Theorem 4.3. This is because, for reversible networks with cycles, it is known that the detailed balancing condition entails further conditions on the rate constants above and beyond complex balancing [14,18]. We can still, however, construct a complex balanced realization from an arbitrary reversible network according to Theorem 4.3. Since detailed balancing implies no further dynamical information above and beyond complex balancing, for all practical purposes this is as far as we need to go.

It is also worth noting that this algorithm cannot determine networks which are linear conjugate to a given structurally-fixed network; it can only find dynamically equivalent networks. This is clear since linear conjugacy requires that

$$Y \cdot A_k = T \cdot Y \cdot A_b \tag{36}$$

where $T = \text{diag}\{\mathbf{c}\}$ and $\mathbf{c} \in \mathbb{R}^m_{>0}$. Regardless of whether we consider the transformation T on the left-hand-side or right-hand-side of (36), it produces a non-linear condition and therefore cannot currently be placed within the existing MILP framework.

4.3 Examples

In this section, we introduce a few examples which illustrate how the algorithm for producing weakly reversible, reversible, and complex balanced realizations of a structurally-fixed but parameter-variable network works.
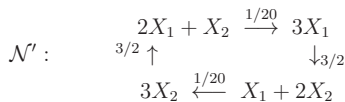
Example 3: Consider the reaction network $\mathcal{N}$ given by

$$\mathcal{N}: \qquad 2X_1 + X_2 \xrightarrow{\alpha} 3X_1 \underset{1}{\overset{1}{\rightleftarrows}} 3X_2 \xleftarrow{\alpha} X_1 + 2X_2$$
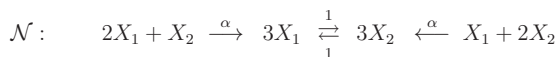
where $\alpha > 0$. Despite the reversible step in the central reaction, $\mathcal{N}$ is neither fully nor weakly reversible and therefore the dynamics do not fall within the scope of the theory of such networks.

We want to check whether there are weakly or fully reversible networks which are dynamically equivalent to $\mathcal{N}$ for some value of α. We set $C_1 = 2X_1 + X_2$, $C_2 = 3X_1$, $C_3 = 3X_2$, and $C_4 = X_1 + 2X_2$. Searching for a sparse network (9) in GLPK over the weakly reversible constraints (30), (31), (33), (13) and (14), with the additional constraints $[A_k]_{21} = [A_k]_{34}$ and $[A_k]_{23} = [A_k]_{32} = 1$ and the bounds $\epsilon = 1/20$ and $u_{ij} = 20$ $i,j =$

$1, \ldots 4$, $i \neq j$, gives the alternative realization

$$\mathcal{N}': \qquad \begin{array}{ccc} 2X_1 + X_2 & \xrightarrow{1/20} & 3X_1 \\ {\scriptstyle 3/2} \uparrow & & \downarrow {\scriptstyle 3/2} \\ 3X_2 & \xleftarrow{1/20} & X_1 + 2X_2 \end{array}$$

and $\alpha = 1/20$ for the original network $\mathcal{N}$. The network $\mathcal{N}'$ has the corresponding kinetics matrix
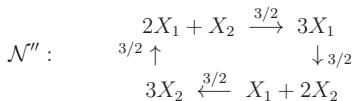
$$A_k' = \begin{bmatrix} -\frac{1}{20} & 0 & \frac{3}{2} & 0 \\ \frac{1}{20} & -\frac{3}{2} & 0 & 0 \\ 0 & 0 & -\frac{3}{2} & \frac{1}{20} \\ 0 & \frac{3}{2} & 0 & -\frac{1}{20} \end{bmatrix}$$

and the positive equilibrium concentration $(x_1^*, x_2^*) = (1,1)$. It can easily be checked that $\mathcal{N}'$ is not complex balanced at this equilibrium concentration.

In order to construct a complex balanced network $\mathcal{N}''$ with the same structure as $\mathcal{N}'$ by (35) we need to determine a vector $\mathbf{b} \in \mathbb{R}_{>0}^4$ such that $\mathbf{b} \in \ker(A_k')$. It can be easily checked that the vector $\mathbf{b} = [3/2 \ 1/20 \ 1/20 \ 3/2]^T$ works; however, using this choice of $\mathbf{b}$ produces a set of rate constants by
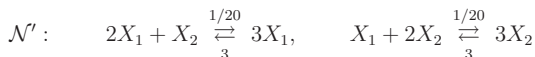
$$A_k \cdot \text{diag}\left\{ \frac{\mathbf{b}}{\Psi(\mathbf{x}^*)} \right\}$$

which clearly violates the condition $[A_k]_{23} = [A_k]_{32} = 1$. This can be solved by choosing another multiple of $\mathbf{b}$. In fact, we can see that for the vector $\mathbf{b} = [30 \ 1 \ 1 \ 30]^T$ the appropriate rate constant choices for $\mathcal{N}$ occur by choosing $\alpha = 3/2$ and the corresponding complex balanced realization given by (35) is

$$\mathcal{N}'': \qquad \begin{array}{ccc} 2X_1 + X_2 & \xrightarrow{3/2} & 3X_1 \\ {\scriptstyle 3/2} \uparrow & & \downarrow {\scriptstyle 3/2} \\ 3X_2 & \xleftarrow{3/2} & X_1 + 2X_2 \end{array}$$

(This corresponds to a scaling by $3/2$ of the 'block' network given in [24]. From the analysis presented in that paper, we know the network is complex balanced only for this particular value of α.)

We may also be interested in fully reversible alternative realizations of $\mathcal{N}$. Replacing the constraints (13) and (14) in the above algorithm by (15) and running in GLPK gives the network

$$\mathcal{N}': \qquad 2X_1 + X_2 \underset{3}{\overset{1/20}{\rightleftarrows}} 3X_1, \qquad X_1 + 2X_2 \underset{3}{\overset{1/20}{\rightleftarrows}} 3X_2$$

and $\alpha = 1/20$ for the network $\mathcal{N}$. The corresponding kinetics matrix is

$$A_k' = \begin{bmatrix} -\frac{1}{20} & 3 & 0 & 0 \\ \frac{1}{20} & -3 & 0 & 0 \\ 0 & 0 & -3 & \frac{1}{20} \\ 0 & 0 & 3 & -\frac{1}{20} \end{bmatrix}$$

which has the positive equilibrium concentration $(x_1^*, x_2^*) = (1,1)$ which is neither complex nor detailed balanced. In order to find a complex balanced network with the same structure as $\mathcal{N}'$ we notice that the kernel of A_k' is given by the span of $[3\ \ 1/20\ \ 0\ \ 0]^T$ and $[0\ \ 0\ \ 1/20\ \ 3]^T$. In order to preserve the property $[A_k]_{23} = [A_k]_{32} = 1$ for the network $\mathcal{N}$ we need to choose $\mathbf{b} = [60\ 1\ 1\ 60]^T$. This gives the value of $\alpha = 3$ for the network $\mathcal{N}$ and the complex balanced realization
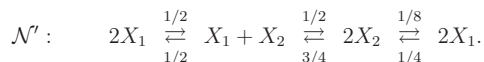
$$\mathcal{N}'' : \qquad 2X_1 + X_2 \underset{3}{\overset{3}{\rightleftarrows}} 3X_1, \qquad X_1 + 2X_2 \underset{3}{\overset{3}{\rightleftarrows}} 3X_2.$$
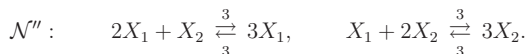
It can easily be checked that $\mathcal{N}''$ is also detailed balanced.

Example 4: Consider the reaction network $\mathcal{N}$ given by

$$\mathcal{N} : \qquad 2X_1 \overset{1}{\longrightarrow} X_1 + X_2 \overset{1}{\longleftarrow} 2X_2$$
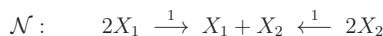
and the reversible alternative realization $\mathcal{N}'$ given by

$$\mathcal{N}' : \qquad 2X_1 \underset{1/2}{\overset{1/2}{\rightleftarrows}} X_1 + X_2 \underset{3/4}{\overset{1/2}{\rightleftarrows}} 2X_2 \underset{1/4}{\overset{1/8}{\rightleftarrows}} 2X_1.$$

If we make the associations $C_1 = 2X_1$, $C_2 = X_1 + X_2$ and $C_3 = 2X_2$, then the network $\mathcal{N}'$ has the corresponding kinetics matrix

$$A_k' = \begin{bmatrix} -\frac{3}{4} & \frac{1}{2} & \frac{1}{8} \\ \frac{1}{2} & -1 & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{2} & -\frac{7}{8} \end{bmatrix}.$$

Choosing any equilibrium concentration $\mathbf{x}^* \in \mathbb{R}_{>0}^2$ we have $A_k' \cdot \Psi(\mathbf{x}^*) \neq \mathbf{0}$ so that the network is not complex balanced.

We wish to construct a complex balanced realization $\mathcal{N}''$ using the methodology outlined in the proof of Theorem 4.3. We choose the equilibrium value $(x_1^*, x_2^*) = (1,1)$ so that $\Psi(\mathbf{x}^*) = [1\ 1\ 1]^T$ and the kernel vector $\mathbf{b} = [1\ 5/4\ 1]^T$. According to (35) we have

$$A_k'' = \begin{bmatrix} -\frac{3}{4} & \frac{1}{2} & \frac{1}{8} \\ \frac{1}{2} & -1 & \frac{3}{4} \\ \frac{1}{4} & \frac{1}{2} & -\frac{7}{8} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{5}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{3}{4} & \frac{5}{8} & \frac{1}{8} \\ \frac{1}{2} & -\frac{5}{4} & \frac{3}{4} \\ \frac{1}{4} & \frac{5}{8} & -\frac{7}{8} \end{bmatrix}.$$

It can be easily checked that the corresponding network $\mathcal{N}''$ is complex balanced and is dynamically equivalent to $\mathcal{N}$. (In general the rate constants of $\mathcal{N}$ may change; however, in this case we have $A_k \cdot \text{diag}\{\mathbf{b}/\Psi(\mathbf{x}^*)\} = A_k$.)

Despite being complex balanced and reversible, the realization $\mathcal{N}''$ is not detailed balanced according to (20). We might wonder if we can construct an alternative realization in a similar manner as (35). We can easily see, however, that applying the detailed

balancing conditions to $A_k \cdot \text{diag}\{\mathbf{c}\}$, where $\mathbf{c} \in \mathbb{R}_{>0}^3$ is arbitrary, produces the unsatisfiable set of conditions

$$\frac{c_1}{2}x_1^2 = \frac{c_2}{2}x_1x_2, \qquad \frac{c_1}{4}x_1^2 = \frac{c_3}{8}x_2^2, \qquad \frac{c_2}{2}x_1x_2 = \frac{3c_3}{4}x_2^2.$$

In other words, we cannot always construct a detailed balanced network from an arbitrary reversible network $\mathcal{N}'$ in the same constructive manner used for complex balanced networks from weakly reversible networks.

5   Conclusions and Future Work

In this paper, new computational methods were presented for finding networks which are linearly conjugate and dynamically equivalent to a given chemical reaction network endowed with mass-action kinetics.

It was demonstrated that finding a dense or sparse reversible, detailed balanced and complex balanced network which is linearly conjugate to a given network can be framed as a MILP optimization problem. The case of determining conjugate networks which have the greatest and fewest number of complexes was also extended to the case of non-trivial linear conjugacies.

It was shown that, similarly to the case of dynamical equivalence [37], the graph structure of linearly conjugate dense networks containing the maximal number of nonzero reaction rate coefficients is unique, and that the unweighted directed reaction graph of any linearly conjugate network is a proper subgraph of the unweighted directed reaction graph of the dense one if the set of complexes is given. Moreover, it was proved that arbitrary equilibrium points of the initial network can be used for the existence checking and computation of linearly conjugate complex balanced and detailed balanced networks.

Additionally, the problem of dynamical equivalence was studied when only the structure of the initial network $\mathcal{N}$ is fixed, but its rate constants can take any positive value. It was shown that in this case, the computation of dense and sparse weakly reversible and reversible networks $\mathcal{N}'$ can also be formulated as a MILP problem. It was also shown that complex balanced networks $\mathcal{N}''$ with identical structure to $\mathcal{N}'$ can be constructed from these realizations and corresponded to an alternative choice of rate constants for $\mathcal{N}$. This modification allows us to scan through a range of parameters as part of the procedure and therefore answer questions about a network based on its structure alone. The operation of the developed methods were illustrated on numerical examples. The achievements further

extend the applicability range of many existing structure-dependent results of chemical reaction network theory.

Further areas of research and open questions include:

1. The procedure introduced to determine structurally dynamically equivalence networks is as of yet unable to search through non-trivial linearly conjugate networks in a linear manner. As such, many networks with potentially insightful information about the dynamics of a given network are being overlooked. Incorporating non-trivial linear conjugacy into a manageable optimization framework is therefore of primary interest.

2. The results obtained so far depend on the reaction networks under consideration having mass-action kinetics (1). Conjugacy and computational results for other widely-used kinetic schemes (e.g. Michaelis-Menten kinetics [29], Hill kinetics [21]) would greatly expand the scope of applicability of these methods.

3. There are many classes of networks with known behaviour lying outside the scope of weakly reversible network theory [4, 8, 9]. Determining constraints which could restrict our search to within these classes of networks would broaden the scope of dynamical behaviours (e.g. periodicity, oscillatory behaviour, multistability, etc.) we could guarantee through this computational procedure.

References

[1] D. Anderson, Global asymptotic stability for a class of nonlinear chemical equations, *SIAM J. Appl. Math.* **68** (2008) 1464–1476.

[2] D. Anderson, A. Shiu, The dynamics of weakly reversible population processes near facets, *SIAM J. Appl. Math.* **70** (2010) 1840–1858.

[3] D. Angeli, P. Leenheer, E. Sontag, A petri net approach to the study of persistence in chemical reaction networks, *Math. Biosci.* **210** (2007) 598–618.

[4] D. Angeli, E. Sontag, Translation-invariant monotone systems, and a global convergence result for enzymatic futile cycles, *Nonlinear Analysis Series B: Real World Applications* **9** (2008) 128–140.

[5] M. Banaji, Monotonicity in chemical reaction systems, *Dyn. Syst.* **24** (2009) 1–30.

[6] M. Banaji, D. Angeli, Convergence in strongly monotone systems with an increasing first integral, *SIAM J. Math. Anal.* **42** (2010) 334–353.

[7] G. Craciun, A. Dickenstein, A. Shiu, B. Sturmfels, Toric dynamical systems, *J. Symbolic Comput.* **44** (2009) 1551–1565.

[8] G. Craciun, M. Feinberg, Multiple equilibria in complex chemical reaction networks: I. the injectivity property, *SIAM J. Appl. Math* **65** (2005) 1526–1546.

[9] G. Craciun, M. Feinberg, Multiple equilibria in complex chemical reaction networks: II. The species–reaction graph, *SIAM J. Appl. Math* **66** (2006) 1321–1338.

[10] G. Craciun, M. Feinberg, Y. Tang, Understanding bistability in complex enzyme–driven reaction networks, *Proc. Natl. Acad. Sci. USA* **103** (2006) 8697–8702.

[11] G. Craciun, C. Pantea, Identifiability of chemical reaction networks, *J. Math Chem.* **44** (2008) 244–259.

[12] G. Craciun, C. Pantea, F. Nazarov, Persistence and permanence of mass–action and power–law dynamical systems, arxiv:1010.3050.

[13] J. Deng, M. Feinberg, C. Jones, A. Nachman, On the steady states of weakly reversible chemical reaction networks, unpublished.

[14] A. Dickenstein, M. Pérez Millán, How far is complex balancing from detailed balancing? *Bull. Math. Biol.* **73** (2011) 811–828.

[15] G. Farkas, Kinetic lumping schemes, *Chem. Eng. Sci.* **54** (1999) 3909–3915.

[16] M. Feinberg, Lectures on chemical reaction networks, Unpublished written versions of lectures given at the Mathematics Research Center, University of Wisconsin.

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

[18] M. Feinberg, Necessary and sufficient conditions for detailed balancing in mass action systems of arbitrary complexity, *Chem. Eng. Sci.* **44** (1989) 1819–1827.

[19] K. Gatermann, B. Huber, A family of sparse polynomial systems arising in chemical reaction systems, *J. Symbolic Comput.* **33** (2002) 275–305.

[20] V. Hárs, J. Tóth, On the inverse problem of reaction kinetics, *Coll. Math. Soc. J. Bolyai* **30** (1981) 363–379.

[21] A. Hill, The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves, *J. Physiol.* **40** (1910) 4–7.

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

[23] F. Horn, The dynamics of open reaction systems: Mathematical aspects of chemical and biochemical problems and quantum chemistry, *SIAM-AMS Proceedings* **8** (1974) 125–137.

[24] F. Horn, R. Jackson, General mass action kinetics, *Arch. Ration. Mech. Anal.* **47** (1972) 187–194.

[25] M. D. Johnston, D. Siegel, Linear conjugacy of chemical reaction networks, *J. Math. Chem.* **49** (2011) 1263–1282.

[26] M. D. Johnston, D. Siegel, Weak dynamic non-emptiability and persistence of chemical kinetics systems, *SIAM J. Appl. Math.* **71** (2011) 1263–1279.

[27] M. D. Johnston, D. Siegel, G. Szederkényi, A linear programming approach to weak reversibility and linear conjugacy of chemical reaction networks, *J. Math. Chem.*, in press.

[28] G. Li, H. Rabitz, A general analysis of exact lumping in chemical kinetics, *Chem. Eng. Sci.* **44** (1989) 1413–1430.

[29] L. Michaelis, M. Menten, Die kinetik der invertinwirkung, *Biochem. Z.* **49** (1913) 333–369.

[30] R. Raman, I. E. Grossmann, Modelling and computational techniques for logic based integer programming, *Comput. Chem. Engin.* **18** (1994) 563–578.

[31] P. Schlosser, M. Feinberg, A theory of multiple steady states in isothermal homogeneous CFSTRs with many reactions, *Chem. Eng. Sci.* **49** (1994) 1749–1767.

[32] D. Siegel, M. D. Johnston, A stratum approach to global stability of complex balanced systems, *Dyn. Syst.* **26** (2011) 125–146.

[33] E. Sontag, Structure and stability of certain chemical networks and applications to the kinetic proofreading model of t-cell receptor signal transduction, *IEEE Trans. Automat. Control* **46** (2001) 1028–1047.

[34] G. Szederkényi, Comment on "identifiability of chemical reaction networks" by G. Craciun and C. Pantea, *J. Math. Chem.* **45** (2009) 1172–1174.

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

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

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

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

[39] J. Tóth, G. Li, H. Rabitz, A. S. Tomlin, The effect of lumping and expanding on kinetic differential equations, *SIAM J. Appl. Math.* **57** (1997) 1531–1556.

[40] J. Wei, J. C. W. Kuo, A lumping analysis in monomolecular reaction systems, *Ind. Eng. Chem. Fundamentals* **8** (1969) 114–133.