MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **93** (2025) 631–687
ISSN: 0340–6253
doi: 10.46793/match.93-3.631V

Weakly Reversible Realizations of Biochemical Systems Theory (BST) Models

Exequiel Jun V. Villejo[a], **Dylan Antonio SJ. Talabis**[b*]

[a] *Institute of Mathematics, University of the Philippines Diliman, Quezon City 1101, Philippines*

[b] *Institute of Mathematical Sciences, University of the Philippines, Los Baños, Laguna 4031, Philippines*

evvillejo@up.edu.ph, dstalabis1@up.edu.ph

(Received September 3, 2024)

Abstract

Biochemical Systems Theory (BST) is a modeling framework that employs power-law formulations to effectively capture the inherent nonlinearities and heterogeneity of biological systems. Recent research has shown that BST models can be modelled by reaction networks. However, many key results in Chemical Reaction Network Theory (CRNT) rely on the condition of weak reversibility - a property often absent in reaction networks derived from BST models. To address this challenge, this paper develops algorithms for constructing weakly reversible realizations of two variants of BST models: S-systems and General Mass Action (GMA) systems. By applying these algorithms, fundamental network properties are simplified, and recent CRNT results regarding the steady states of such systems are validated. Additionally, some of these algorithms yield deficiency zero networks - a necessary property for the existence of complex-balanced steady states. Finally, the proposed algorithms are applied to the GMA representation of the carbon cycle models by Anderies et al. and Heck et al., demonstrating the existence of concentration robustness in these models.

*Corresponding author.

1 Introduction

For over 150 years, the mass action law, which was first formulated by Cato Maximillian Guldberg and Peter Waage in 1864 and later clarified by Jacobus van 't Hoff in 1877, has established itself as the most used default model in mathematical modelling to the point that it is considered an undisputed truth [11, 29]. However, this formulation quickly fails in realistic situations such that even simple chemical reaction systems lead to nonlinearities [10,11]. It is for this reason that Michael Savageau proposed a new framework he called Biochemical Systems Theory (BST) in 1969 to take into account nature's heterogeneity [11]. This paved the way for the analysis of dynamic and large-scale biological and biochemical systems. Instead of using mass action law, Savageau used power law formulations, which can be thought of as a direct generalization of Guldberg and Waage's mass action law, as core representations for such processes [10, 11].

It has been found out in [3] that BST models can be represented by chemical reaction networks, which means that results in Chemical Reaction Network Theory (CRNT) can be applied to analyze BST models. Now, many results in CRNT are hinged on the condition that reaction networks be weakly reversible (i.e., all components of a linkage class are strongly connected with each other). For example, the deficiency theorems require that the reaction networks be weakly reversible so that conclusions can be made on the dynamics of the associated system, such as on the non-emptiness of its set of positive steady states [5, 7, 30, 31]. Furthermore, non-weakly reversible systems with zero deficiency are generally undesirable because they cannot admit a positive equilibrium, nor can they support a cyclic composition trajectory in which all species concentrations remain positive [21]. However, weak reversibility is usually not the case for chemical reaction networks underlying BST models [3]. In fact, all of the seventeen BST models of biological systems examined in [3] and [4] were found to be not weakly reversible. This leads to the following question: Is it possible to construct weakly reversible realizations of chemical reaction networks underlying BST models?

Attempts were made to construct weakly reversible realizations of reac-

tion networks. However, the reaction networks considered are only those that are endowed with mass action kinetics. Szederkenyi, Hangos, and Tuza proposed a method under a Mixed Integer Linear Programming (MILP) framework based on elementary graph theory to construct a dense weakly reversible realization [14]. Here, the method requires potentially multiple MILP optimizations which are known to be NP-hard [22]. As a remedy, the weak reversibility part is reformulated as a linear constraint which is based on the fact that a reaction graph corresponding to a Laplacian matrix A_k is weakly reversible if and only if there exists a vector $b \in \mathbb{R}^m_{>0} \cap \ker(A_k)$ where m is the number of species in the reaction network [22]. In this reformulation, determining a weakly reversible realization can now be done in a single MILP step.

As mentioned, these attempts of finding a dynamically equivalent weakly reversible realization are valid only under the assumption of mass action kinetics which does not answer the original question posed above.

This paper presents algorithms for constructing weakly reversible realizations of reaction networks underlying BST models. In particular, it considers two variants of BST models, namely General Mass Action (GMA) systems and S-systems. For S-systems, we develop an algorithm that will preserve the reaction vectors of the influx and efflux reactions of the algorithm in [3]. Notably, this new dynamically equivalent network is weakly reversible and deficiency zero.

For GMA systems, we consider two approaches. For small scale systems, we use the characterization of the incidence matrix I_a that encodes a weakly reversible network. Here, we determine the positive entries of the molecularity matrix Y by solving the induced linear system from the matrix equation $Y \cdot I_a = N$, where I_a is selected to ensure that the associated network is weakly reversible. To narrow down the candidates for I_a in this formulation, we use the result in [8, 15] which states that if the reaction network is positive-dependent, then the network has a weakly reversible realization. We take advantage of this by breaking down the network into positive-dependent subnetworks and focus on finding weakly reversible realizations for each subnetwork. For large scale GMA systems, we also consider poly-PL kinetics, introduced in [6], to apply the S-system algo-

rithm. These poly-PL kinetics, i.e. positive linear combinations of power law kinetics, naturally occur in the kinetic system realizations of GMA systems. Because the resulting networks are weakly reversible and deficiency zero, we can leverage the power-law results of Talabis et al. [5–7] and the parametrization results of Müller and Regensburger [30]. Lastly, we apply the algorithms proposed in this paper for concentration robustness analysis. We demonstrate the effectiveness of our approach by considering the carbon cycle models Anderies et al. [17] and Heck et al. [32], showcasing the applicability of our methods in analyzing such models.

The paper is organized as follows: Section 2 presents some basic notions on CRNT and BST models as needed in the succeeding sections. In Section 3, we present in detail a characterization of the structure of an I_a that encodes weakly reversible networks. This section also discusses an approach to be more efficient in choosing an I_a in the implementation of the algorithm. Section 4 offers a complete, step-by-step description of the algorithms for finding a weakly reversible realization of BST models. Additionally, various network and kinetic properties of the constructed weakly reversible realizations are analyzed and compared to those of the reaction network representations in [3]. In Section 5, we apply our algorithms to identify key features of the system, such as concentration robustness. The results are illustrated using the GMA representation of the carbon cycle models by Anderies et al [17] and Heck et al [32]. Lastly, an overall summary is provided in Section 6.

2 Preliminaries

2.1 Fundamentals of reaction networks and kinetic systems

A chemical reaction network (CRN) is defined by three sets. First, we have the **set of species** which is denoted by $\mathscr{S} = X_1, X_2, \ldots, X_m$. Its cardinality is equal to m. Second, we have the **set of complexes**, denoted by $\mathscr{C} = C_1, C_2, \ldots, C_n$, which are linear combinations of the species with nonnegative stoichiometric coefficients. Its cardinality is equal to n.

Lastly, we have the **set of reactions** $\mathscr{R}$ which are ordered pairs of distinct complexes. The ordered pair $(C_i, C_j) \in \mathscr{R}$ is denoted in the CRN as $C_i \to C_j$ with a nonnegative weight, k_{ij}, assigned to it called the **reaction rate coefficient**. Here, C_i is called the **reactant complex** while C_j is the **product complex**. The cardinality of $\mathscr{R}$ is equal to r.

A CRN can be represented by a directed graph. Recall that a directed graph (or digraph) D consists of a non-empty finite set $V(D)$ of elements called vertices and a finite set $E(D)$ of ordered pairs of distinct vertices called edges. In CRN, the set of vertices in the digraph is $\mathscr{C}$ while the set of edges is $\mathscr{R}$. A directed walk is a sequence of edges $\{E_1, E_2, ..., E_{n-1}\}$ connecting a sequence of vertices $\{V_1, V_2, ..., V_n\}$ such that $E_i = (V_i, V_{i+1})$ for $i = 1, ..., n-1$. A directed path is a directed walk with distinct edges and vertices. Two vertices are said to be connected if there is a directed path between them. Furthermore, two vertices are strongly connected if there is a directed path to and from each other.

Given these terminologies, we say that two complexes C_i and C_j are **connected** if we can find a series of reactions that connects C_i to C_j, or from C_j to C_i. On the other hand, they are **strongly connected** if we can find a series of reactions that connects C_i to C_j, and vice versa.

A set of complexes is called a **linkage class** of the CRN, denoted by $\mathscr{L}^i$, if the complexes in the set are connected but not to any other complex that is not in the set. Moreover, it is a **strong linkage class** if the complexes are strongly connected. It can also be a subset of the linkage class where any two complexes are connected by a directed path in each direction. In other words, a (strong) linkage class is a maximal set of (strongly) connected complexes. A complex, if not strongly connected to other complex/es in the CRN, is a **trivial strong linkage class**. We denote the number of linkage classes by l and the number of strong linkage classes by sl. A CRN is **weakly reversible** if $l = sl$ (i.e., every linkage class is a strong linkage class).

Running Example 1 - Part 1. To illustrate the concepts above, consider the chemical reaction network below.

$$X_1 \xrightarrow{k_1} 2X_1 + X_2$$

$$0 \underset{k_3}{\overset{k_2}{\rightleftarrows}} X_3 \xrightarrow{k_4} X_4 + X_5$$

The k_i's are called the reaction rate coefficients. We have $m = 5$ (species), $n = 5$ (complexes), and $r = 4$ (reactions), where

$$\mathscr{S} = \{X_1, X_2, X_3, X_4, X_5\}, \quad \mathscr{C} = \{X_1, 2X_1 + X_2, 0, X_3, X_4 + X_5\},$$

$$\mathscr{R} = \{(X_1, 2X_1 + X_2), (0, X_3), (X_3, 0), (X_3, X_4 + X_5)\}$$

The entity "0" in the CRN is called the zero complex which is just the zero vector in $\mathbb{R}^{\mathscr{S}}$. The zero complex helps model the creation or annihilation of species in the network. The zero complex and the complex $X_4 + X_5$ are connected since we can find a series of reactions from 0 to $X_4 + X_5$. However, they are not strongly connected. An example of strongly connected complexes are the zero complex and X_3. Here, there are two linkage classes, namely $\mathscr{L}^1 = \{X_1, 2X_1 + X_2\}$ and $\mathscr{L}^2 = \{0, X_3, X_4 + X_5\}$. But there are 4 strong linkage classes, namely $\{X_1\}, \{2X_1 + X_2\}, \{0, X_3\}$,and $\{X_4 + X_5\}$. So, $l = 2$ but $sl = 4$. Therefore, the CRN $\mathscr{N}$ is not weakly reversible.

The matrices that will be defined characterize a CRN. The **molecularity matrix** Y is an $m \times n$ matrix where the (i, j)-th entry is the stoichiometric coefficient of the species X_i in complex C_j. The **incidence matrix** I_a is an $n \times r$ matrix where each row corresponds to a complex and each column corresponds to a reaction, satisfying

$$(I_a)_{(i,j)} = \begin{cases} -1 & \text{if } i \text{ is the reactant complex of reaction } j \in \mathscr{R}, \\ 1 & \text{if } i \text{ is the product complex of reaction } j \in \mathscr{R}, \\ 0 & \text{otherwise.} \end{cases}$$

Lastly, the **stoichiometric matrix** N is an $m \times r$ matrix whose entries are taken from the reaction vectors. The reaction vectors of the CRN are

the members of the set $\{C_j - C_i \in \mathbb{R}^m | C_i \to C_j \in \mathscr{R}\}$. The span of this set is called the **stoichiometric subspace**. Alternatively, $N = Y \cdot I_a$.

Running Example 1 - Part 2. In Running Example 1, the matrices Y, I_a, and N are

$$Y = \begin{array}{c} \begin{matrix} X_1 & 2X_1 + X_2 & 0 & X_3 & X_4 + X_5 \end{matrix} \\ \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{matrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{matrix}$$

$$I_a = \begin{array}{c} \begin{matrix} R_1 & R_2 & R_3 & R_4 \end{matrix} \\ \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{matrix} X_1 \\ 2X_1 + X_2 \\ 0 \\ X_3 \\ X_4 + X_5 \end{matrix} \qquad N = \begin{array}{c} \begin{matrix} R_1 & R_2 & R_3 & R_4 \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{matrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{matrix}$$

Here, we denote the reactions by R_i such that $R_1 : X_1 \to 2X_1 + X_2$, $R_2 : 0 \to X_3$, $R_3 : X_3 \to 0$, and $R_4 : X_3 \to X_4 + X_5$.

A kinetics K of a CRN is defined as an assignment to each reaction $C_i \to C_j \in \mathscr{R}$ of a continuously differentiable rate function $K_{C_i \to C_j} : \bar{\mathbb{R}}_+^{\mathscr{S}} \to \bar{\mathbb{R}}_+$ such that if $K_{C_i \to C_j}(c) > 0$ then $\text{supp}(C_i) \subset \text{supp}(c)$. $\text{Supp}(C_i)$ are the species that appear in C_i. The pair $(\mathscr{N}, K)$ is called a chemical kinetic system (CKS).

Under **power law kinetics** (PLK), the rate at which a reaction occurs is given by

$$\mathscr{K}_i(x) = k_i \prod_{j=1}^{m} x_j^{F_{ij}}, \qquad \forall i \in \{1, 2, \ldots, r\}$$

with rate constants $k_i > 0$ and $F_{ij} \in \mathbb{R}$ known as **kinetic orders** defined in the $r \times m$ **kinetic order matrix** F. PLK offers more flexibility in modelling systems in biochemistry, epidemics, etc. [9]. A special class

of PLK is the **power law reactant-determined kinetics (PL-RDK)** where reactions with identical reactant complexes have the same kinetics orders. If this condition is not met, the kinetic system falls under **PL-NDK**.

Running Example 1 - Part 3. In Running Example 1, we can define the kinetics of the CRN under PLK as follows

$$K(x) = \begin{bmatrix} k_1 X_1^{f_{11}} \\ k_2 \\ k_3 X_3^{f_{33}} \\ k_4 X_3^{f_{43}} \end{bmatrix}$$

where $f_{ij} \in F$. The dynamical system of the CRN of our running example can be written as

$$\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \\ \dot{X}_3 \\ \dot{X}_4 \end{bmatrix} = \begin{array}{c} \begin{matrix} R_1 & R_2 & R_3 & R_4 \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{bmatrix} k_1 X_1^{f_{11}} \\ k_2 \\ k_3 X_3^{f_{33}} \\ k_4 X_3^{f_{43}} \end{bmatrix} = N \cdot K(x).$$

It is possible for different CRNs to have the same set of ordinary differential equations. This motivates the concept of dynamical equivalence. Two CRNs are said to be **dynamically equivalent** if they give rise to the same set of ordinary differential equations. In such case, we say that these CRNs are **realizations** of the associated dynamical system.

Another type of kinetics is the **Poly-PL kinetics** (PYK) which are kinetic systems consisting of nonnegative linear combinations of power law functions. Similar to PLK, the domain of PYK is $\mathbb{R}^m_>$. Clearly, both PLK and PYK generate the same species formation rate which is given by the power law dynamical systems. We can define reactant-determined kinetics for PYK (PY-RDK) similarly as PL-RDK. In 2019, Talabis et al. [6] formally defined PYK as follows.

Definition 1. A kinetics $K : \mathbb{R}^m_> \to \mathbb{R}^r$ is a **poly-PL kinetics** if

$$K_i(x) = k_i(a_{i,1}x^{F_{i,1}} + ... + a_{i,j}x^{F_{i,j}}) \quad \forall i \in \{1, 2, ..., r\}$$

written in lexicographic order with $k_i \in \mathbb{R}_+$, $F_{i,j}, a_{i,j} \in \mathbb{R}^m$ and $j \in \{1, 2, ..., h_i\}$ (where h_i is the number of terms in reaction i). Poly-PL kinetics is defined by $r \times m$ matrices $F_{i,k} = (F_{ij})$, called the **kinetic order matrices** while vectors $k = (k_i)$ and $(a_{i,\cdot}) \in \mathbb{R}^r_>$ are called the **rate vector** and **poly-rate vectors**, respectively.

In 2018, Talabis et al. [5] defined the concepts T and $\hat{T}$-matrices as follows.

Definition 2. The $m \times n_r$ **T-matrix** is the truncated $\tilde{Y}$ where the non-reactant columns are deleted and n_r is the number of reactant complexes. The $\hat{T}$ **matrix** is constructed from the T matrix such that

$$\hat{T} = \begin{bmatrix} T \\ L^T \end{bmatrix}$$

where L is the $n_r \times l$ matrix defined by $L = \begin{bmatrix} e_1, e_2, \ldots, e_l \end{bmatrix}$ where e^i is a characteristic vector for linkage class $\mathscr{L}^i$.

It is also in [5] where Talabis et al. defined PL-TIK systems which is a subclass of PL-RDK systems.

Definition 3. A **PL-TIK** kinetics is $\hat{T}$-rank maximal if the column rank of $\hat{T}$ is maximal.

Alternatively, we can determine if a system is PL-TIK by looking at the concept of the **kinetic order deficiency**. It measures the degree of kinetic interactions of the PL-RDK system. The kinetic order deficiency $\hat{\delta}$ can be computed by $\hat{\delta} = n_r - \hat{q}$ where $\hat{q}$=rank($\hat{T}$). If a system is PL-TIK, it follows that rank($\hat{T}$) $= n_r$. Hence, a system is PL-TIK if and only if $\hat{\delta} = 0$.

There is an analog of Definitions 2 and 3 for PYK systems which can be found in [6]. In 2019, Talabis et al. [6] defined the concept of T_κ and $\hat{T}_\kappa$ matrices, and a PY-TIK system as follows.

Definition 4. The $m \times n_r$ **poly T-matrix** T_κ $(\forall \kappa \in \{1, 2, ..., h\})$ is the truncated $\widetilde{Y}_\kappa$ where the non-reactant columns are deleted. Define the $n_r \times l$ matrix $L = (e_1, e_2, ..., e_l)$ where e^i is a characteristic vector for linkage class $\mathscr{L}^i$. The block matrix $\widehat{T}_\kappa \in \mathbb{R}^{(m+l)\times n_r}$ $(\forall \kappa \in \{1, 2, ..., h\})$ is defined as

$$\hat{T}_\kappa = \begin{pmatrix} T_\kappa \\ L^T \end{pmatrix}$$

Definition 5. A PYK system is **PY-TIK** if every $\hat{T}_\kappa$ is rank maximal (i.e., column rank of $\hat{T}_\kappa$ is maximal).

2.2 Equilibria and network decomposition

In this subsection, we review some definitions and earlier results on equilibria. Later in this subsection, we will look into the theory of network decomposition.

Once a kinetics is associated to a CRN, we can now determine the rate at which the concentration of each species evolves through time.

Definition 6. The **species rate formation function** $f : \mathbb{R}^m_{\geq 0} \to \mathbb{R}^n$ of a CKS is given by

$$f(x) = N \cdot K(x) = \sum_{C_i \to C_j \in \mathscr{R}} K_{C_i \to C_j}(x)(C_j - C_j)$$

for all $x \in \mathbb{R}^m_{\geq 0}$. The ODE or dynamical system of the CKS is defined as $\dfrac{dx}{dt} = f(x)$. An element c^* of $\mathbb{R}^m_{>0}$ for which $f(c^*) = 0$ is called a **positive equilibrium or steady state**. The **set of positive steady states**, denoted as E_+, is defined by $E_+ = \{x \in \mathbb{R}^m_{>0} | f(x) = 0\}$.

We also have the concept of a complex formation rate function which is the analog of the species formation rate function for complexes.

Definition 7. The **complex formation rate function** $g : \mathbb{R}^m_{>0} \to \mathbb{R}^n$ of a CKS is given by

$$g(x) = I_a \cdot K(x)$$

A **complex balanced steady state** c happens if $g(c) = 0$. A CKS is **complex balanced** if it has a complex balanced steady state. The **set of complex balanced equilibria**, denoted by Z_+, is given by $Z_+ = \{x \in \mathbb{R}^m_{>0} | g(x) = 0\}$.

Complex balanced systems played an important role in the development of the theory of chemical reaction networks. It was Horn and Jackson [12] who first introduced the concept of complex balancing. From the definition above, we can say that a system is complex balanced at a state if for each complex, formation and degradation are at equilibrium. Lastly, it is worth noting that $Z_+(\mathscr{N}, K) \subseteq E_+(\mathscr{N}, K)$.

Positive equilibria and complex balanced equilibria of certain classes of power law systems were explored in [5, 7]. The following results are called deficiency theorems because of the condition that the deficiency be 0 or 1.

Theorem 1 ([5]). *Let $(\mathscr{N}, K)$ be a PL-TIK system with $\delta = 0$. Then $E_+(\mathscr{N}, K) \neq 0$ if and only if $\mathscr{N}$ is weakly reversible. Furthermore, the following also hold:*

i. If $E_+(\mathscr{N}, K) \neq \emptyset$ and $x^* \in E_+(\mathscr{N}, K)$, then

$$E_+(\mathscr{N}, K) = \left\{ x \in \mathbb{R}^m_+ | \log(x) - \log(x^*) \in \left(\tilde{S}_R\right)^\perp \right\}$$

ii. If $E_+(\mathscr{N}, K) \neq \emptyset$, then $|E_+(\mathscr{N}, K) \cap \boldsymbol{Q}| = 1$ for each positive kinetic class $\boldsymbol{Q}$.

Theorem 2 ([5]). *Let $(\mathscr{N}, K)$ be a PL-TIK system satisfying*

i. $\delta_\theta \leq 1, \qquad \theta = 1, 2, \ldots, l$

ii. $\sum_{\theta=1}^{l} \delta_\theta = \delta$

If $\mathscr{N}$ is weakly reversible, then $E_+(\mathscr{N}, K) \neq 0$.

Items i.) and ii.) of Theorem 1 also hold for Theorem 2. The theorems above show that the set of positive equilibria is non-empty if the network is weakly reversible with deficiency equal to either 0 or 1. In Talabis et al [7],

the condition that the deficiency be either zero or one is removed while yielding conclusions about the complex-balanced equilibria. This can be referred as a high deficiency theorem for PL-TIK systems.

Theorem 3 ([7]). *Let $(\mathscr{N}, K)$ be a PL-TIK system. Then $Z_+(\mathscr{N}, K) \neq 0$ if and only if $\mathscr{N}$ is weakly reversible. Furthermore, the following also hold:*

i. If $Z_+(\mathscr{N}, K) \neq 0$ and $x^* \in Z_+(\mathscr{N}, K) \neq 0$, then

$$Z_+(\mathscr{N}, K) = \left\{ x \in \mathbb{R}^m_{\geq} | \log(x) - \log(x^*) \in \left(\tilde{S}_R\right)^{\perp} \right\}$$

ii. If $Z_+(\mathscr{N}, K) \neq 0$, then $|Z_+(\mathscr{N}, K) \cap \boldsymbol{Q}| = 1$ for each positive kinetic reactant flux class $\boldsymbol{Q}$.

We also recall from [6] statements on the existence, parametrization, and uniqueness of complex balanced equilibria for PY-TIK systems:

Theorem 4. *[6] Let $(\mathscr{N}, K)$ be a PY-TIK systems. Then $\mathscr{N}$ is weakly reversible if and only if $Z_+(\mathscr{N}, K) \neq \emptyset$.*

Theorem 5. *[6] Let $(\mathscr{N}, K)$ a weakly reversible poly-PL kinetic system with poly T-matrices T_1, ..., T_h. Consider an arbitrary poly T-matrix T_k.*

(i) if $Z_+(\mathscr{N}, K) \neq \emptyset$ and $x^* \in Z_+(\mathscr{N}, K)$ then

$$Z_+(\mathscr{N}, K) = \left\{ x \in \mathbb{R}^m_{\geq} \middle| \log(x) - \log(x^*) \in (\widetilde{S}_j)^{\perp} \right\}.$$

(ii) if $Z_+(\mathscr{N}, K) \neq \emptyset$ then $|Z_+(\mathscr{N}, K) \cap Q_j| = 1$ for each positive kinetic reactant flux class Q_j.

We now proceed to reviewing the theory of network decomposition by stating the following definitions.

Definition 8. A **decomposition** of $\mathscr{N}$ is a set of subnetworks of $\mathscr{N}$ $(\{\mathscr{N}_1, \mathscr{N}_2, \ldots, \mathscr{N}_k\})$ induced by a partition $\{\mathscr{R}_1, \mathscr{R}_2, \ldots, \mathscr{R}_k\}$ of its reaction set $\mathscr{R}$.

Definition 9. A network decomposition $\mathscr{N} = \mathscr{N}_1 \cup \mathscr{N}_2 \cup \ldots \cup \mathscr{N}_k$ is **independent** if its stoichiometric subspace is a direct sum of the subnetwork stoichiometric subspaces. It is **incidence independent** if the image of the network's incidence map is a direct sum of the images of the incidence maps of the subnetworks.

In Fortun et al [26], it was shown that for an independent decomposition, $\delta \leq \delta_1 + \delta_2 + \ldots + \delta_k$. In Farinas et al [16], it was shown that for an incidence independent decomposition, $\delta \geq \delta_1 + \delta_2 + \ldots + \delta_k$.

Feinberg identified a fundamental relationship between an independent decomposition and the set of positive equilibria of a network.

Theorem 6 ([19, 20])**.** *Let $P(\mathscr{R}) = \{\mathscr{R}_1, \mathscr{R}_2, ..., \mathscr{R}_k\}$ be a partition of a CRN $\mathscr{N}$ and let K be a kinetics on $\mathscr{N}$. If $\mathscr{N} = \mathscr{N}_1 \cup \mathscr{N}_2 \cup ... \cup \mathscr{N}_k$ is the network decomposition of $P(\mathscr{R})$ and $E_+(\mathscr{N}_i, K_i) = \left\{x \in \mathbb{R}^{\mathscr{S}}_{>0} | N_i K_i(x) = 0\right\}$ then*

$$E_+(\mathscr{N}_1, K_1) \cap E_+(\mathscr{N}_2, K_2) \cap ... \cap E_+(\mathscr{N}_k, K_k) \subseteq E_+(\mathscr{N}, K).$$

If the network decomposition is independent, then equality holds.

The following theorem is an analogue of Feinberg's 1987 result for incidence independent decompositions and complex balanced equilibria.

Theorem 7 ([16])**.** *Let $\mathscr{N} = (\mathscr{S}, \mathscr{C}, \mathscr{R})$ be a a CRN and $\mathscr{N}_i = (\mathscr{S}_i, \mathscr{C}_i, \mathscr{R}_i)$ for $i = 1, 2, ..., k$ be the subnetworks of a decomposition. Let K be any kinetics, and $Z_+(\mathscr{N}, K)$ and $Z_+(\mathscr{N}_i, K_i)$ be the sets of complex balanced equilibria of $\mathscr{N}$ and $\mathscr{N}_i$, respectively. Then*

i. $Z_+(\mathscr{N}_1, K_1) \cap Z_+(\mathscr{N}_2, K_2) \cap ... \cap Z_+(\mathscr{N}_k, K_k) \subseteq Z_+(\mathscr{N}, K)$.
If the decomposition is incidence independent, then

ii. $Z_+(\mathscr{N}, K) = Z_+(\mathscr{N}_1, K_1) \cap Z_+(\mathscr{N}_2, K_2) \cap ... \cap Z_+(\mathscr{N}_k, K_k)$, *and*

iii. $Z_+(\mathscr{N}, K) \neq \varnothing$ *implies* $Z_+(\mathscr{N}_i, K_i) \neq \varnothing$ *for each* $i = 1, ..., k$.

2.3 BST models and their CRN representations

There are different variants within BST depending on the rules in setting up the equations for the model such as in the order of flux aggregation

and power law approximation [10]. The two most common variants are the GMA system and S-system.

If every process/reaction is modeled with its own power law representation, we have a **GMA system**. Its general format is as follows

$$\dot{X}_i = \sum_{k=1}^{T_i} \pm\gamma_{ik} \prod_{j=1}^{n+m} X_j^{f_{ikj}}, \qquad i = 1, 2, \ldots, n$$

where n and m are the number of dependent and independent variables, respectively, and T_i is the number of reactions/processes associated to species X_i, γ_j are the rate constants, and f_l are the kinetic orders. In BST, a species is called dependent if it varies with time (i.e., $\dot{X}_i \neq 0$) and independent otherwise. In this formulation, incoming reactions take positive values while outgoing reactions take negative values. On the other hand, if incoming processes or reactions and outgoing processes or reactions are first aggregated and collectively modeled with only one power law representation each, we have an **S-system** [10]. Its general format is as follows

$$\dot{X}_i = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}}, \qquad i = 1, 2, \ldots, n.$$

where α_i and β_i are the rate constants, and g_{ij} and h_{ij} are the kinetic orders. For more details on BST model construction and design, you can refer to [10].

GMA system is more intuitive since it focuses on fluxes [10]. One flux is modeled using one power law term. On the other hand, S-system is simpler because it only involves two terms. It was also shown that S-systems are more accurate representation for functions that start at a small value and monotonically grow toward saturation [10]. Furthermore, GMA systems require more complicated solution methods whereas S-system can only be computed using linear algebra techniques [10].

In [3], a method is developed to represent a BST model as a CRN. For the representation, a biochemical map is used which is like a reaction graph. A directed mass transfer arrow in the biochemical map represents a

reaction in the CRN. Furthermore, regulatory arrows - often represented as broken lines in the biochemical map - are also accounted for which change the connectivity of the CRN representation. These regulatory arrows point from an element in the biochemical map to a reaction which represent the role of that element as an inhibitor or promoter of that particular reaction. If the sign associated to a regulatory arrow is negative, the element from which the arrow originates acts as an inhibitor to the reaction for which the regulatory arrow is pointing to. If the sign is positive, then that element acts as a promoter to the reaction.

Running Example 2 - Part 1. Consider the map in Figure 1.

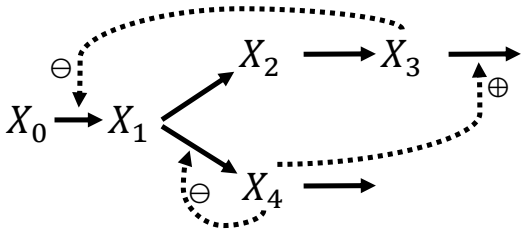


Figure 1. Biochemical Map of Generic pathway with one activating and two inhibitory signals (Figure 3 from [11]).

For the GMA representation, we consider each reaction and regulatory process separately and construct their respective power law formulation. Thus,

$$\begin{aligned}
\dot{X}_1 &= \alpha_1 X_3^{g_{14}} X_0^{g_{11}} - \beta_{11} X_1^{h_{1,11}} - \beta_{12} X_1^{h_{1,12}} X_4^{h_{1,15}} \\
\dot{X}_2 &= \alpha_2 X_1^{g_{22}} - \beta_2 X_2^{h_{23}} \\
\dot{X}_3 &= \alpha_3 X_2^{g_{33}} - \beta_3 X_3^{h_{34}} X_4^{h_{35}} \\
\dot{X}_4 &= \alpha_4 X_1^{g_{42}} X_4^{g_{45}} - \beta_4 X_4^{h_{45}}
\end{aligned}$$

For the S-system representation, we consider all the species that is involved in both the incoming and outgoing reactions, and in the regulatory processes. We then aggregate them each by one term in the equation.

Thus,

$$\begin{aligned}
\dot{X}_1 &= \kappa_1 X_0^{g_{11}} X_3^{g_{41}} - \kappa_2 X_1^{g_{22}} X_4^{g_{52}} \\
\dot{X}_2 &= \kappa_3 X_1^{g_{23}} - \kappa_4 X_2^{g_{34}} \\
\dot{X}_3 &= \kappa_5 X_2^{g_{35}} - \kappa_6 X_3^{g_{46}} X_4^{g_{56}} \\
\dot{X}_4 &= \kappa_7 X_1^{g_{27}} X_4^{g_{57}} - \kappa_8 X_4^{g_{58}}
\end{aligned}$$

As mentioned, it has been found out that BST models can be represented by reaction networks [3]. The goal is to associate a CRN representation such that it will induce the same ODE system as the BST model using its biochemical map representation. In [3], Arceo et al. present an algorithm for transforming Biochemical Systems Theory (BST) models into reaction networks, which we refer to as the BST Algorithm.

Algorithm 1 (BST Algorithm)**.** *A. For GMA systems, we consider first the inflow (i.e., reactions of the form* $0 \to X_i$*) and outflow (i.e., reactions of the form* $X_i \to 0$*) arrows and those without regulatory arrows. In the CRN representation, we include them as they are (e.g.,* $X_1 \to X_2$ *and* $0 \to X_0$*). For those with regulatory arrows (dashed arrows), the rule is that for each interaction* $X_i \to X_j$ *with a regulatory arrow from each* X_k*, we associate the reaction* $X_i + \sum X_k \to X_j + \sum X_k$*.*

B. For S-system, for each dependent variable X_i*, we associate the reaction* $\sum X_{g,j} \to X_i + \sum X_{g,j}$ *for the production (influx) term of* $\dot{X}_i$ *where* $X_{g,j}$ *is a variable with* $g_{ij} \neq 0$ $(i \neq j)$ *in the production (influx) term in the ODE model. Furthermore, we associate the reaction* $X_i + \sum X_{h,j} \to \sum X_{h,j}$ *for the degradation (efflux) term of* $\dot{X}_i$ *where* $X_{h,j}$ *is a variable with* $h_{ij} \neq 0$ $(i \neq j)$ *in the degradation (efflux) term in the ODE model.*

Running Example 2 - Part 2. Consider Running Example 2. Using the BST Algorithm, a realization of the Generalized Mass Action (GMA) model is given by

$$R_1 : 0 \to X_0 \qquad R_5 : X_1 \to X_2$$
$$R_2 : X_0 + X_3 \to X_1 + X_3 \qquad R_6 : X_3 + X_4 \to X_4$$
$$R_3 : X_4 \to 0 \qquad R_7 : X_2 \to X_3$$
$$R_4 : X_1 + X_4 \to 2X_4 \qquad .$$

Meanwhile, the CRN representation of the S-system model is given by

$$R_1 : X_0 + X_3 \to X_0 + X_3 + X_1 \qquad R_5 : X_1 \to X_1 + X_2$$
$$R_2 : X_2 \to X_2 + X_3 \qquad R_6 : X_1 + X_4 \to X_1 + 2X_4$$
$$R_3 : X_1 + X_4 \to X_4 \qquad R_7 : X_2 \to 0$$
$$R_4 : X_3 + X_4 \to X_4 \qquad R_8 : X_4 \to 0$$

3 An I_a that encodes a weakly reversible network

3.1 Characterization of an I_a of a weakly reversible network

In Section 2, it was mentioned that the stoichiometric matrix N is one of the matrices that characterizes a CRN. Every row of N describes all the reactions in which a particular species participates, therefore telling us how the reactions are interconnected [1]. It is also essential in calculating the flux production and degradation of each species of the network [27]. Since we have $\dot{x} = N \cdot K(x)$, the stoichiometric matrix N determines the dynamics of the network.

Now, we wish to construct a weakly reversible network that is dynamically equivalent to the CRN representation of BST models based on the method in Arceo at al [3]. Hence, it is important to keep the stoichiometric matrix as is to preserve the dynamics of the initial network.

Recall that $N = Y \cdot I_a$ where Y is the molecularity matrix and I_a is the incidence matrix. The matrix Y contains the stoichiometric coefficient of the species X_i in complex C_j while the matrix I_a shows how the

complexes are interconnected in the reactions of the network. Since weak reversibility has also something to do with the connectivity of the network, the approach is to construct a network with I_a that encodes a weakly reversible network with the same N as the initial network. The succeeding part of this section determines the structure of an I_a that encodes a weakly reversible network.

Recall that a network is weakly reversible if for every pair of complexes in a linkage class we can find a directed path from and to each other. If that is the case, the minimum number of reactions for a weakly reversible network must be 2. The first one is the reaction that goes from a complex C_i to another complex C_j while the second one is the reaction that goes the opposite direction. So, ***an I_a that encodes a weakly reversible network must contain at least 2 columns***. Another direct implication from the definition of weak reversibility is that each complex in the network must both be a reactant complex and a product complex. ***This corresponds to I_a having both -1 and 1 in each row*** as per definition of the incidence matrix. Consider the network with its I_a given below.

$$X_1 \xrightarrow{k_1} X_2$$

$$X_2 + X_3 \xrightarrow{k_3} X_3 \xleftarrow[k_2]{} 0$$

$$2X_3 \xrightarrow{k_4} X_1 + X_3$$

$$I_a = \begin{array}{c} \begin{matrix} R_1 & R_2 & R_3 & R_4 \end{matrix} \\ \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{matrix} X_1 \\ X_2 \\ X_2 + X_3 \\ X_3 \\ 0 \\ 2X_3 \\ X_1 + X_3 \end{matrix}$$

The network is not weakly reversible. Here, X_2, X_3, and $X_1 + X_3$ are

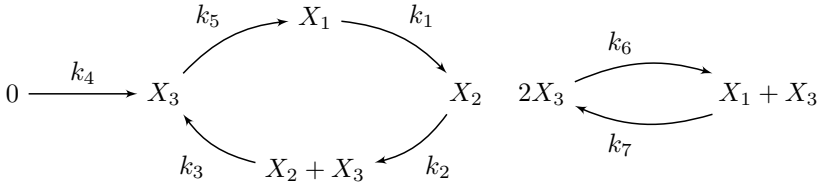
not reactant complexes which correspond to not having -1 in their rows in I_a. Suppose we make them reactant complexes by adding the following reactions below.

$$X_2 \longrightarrow X_2 + X_3, \qquad X_3 \longrightarrow X_1, \qquad X_1 + X_3 \longrightarrow 2X_3$$

We get the updated I_a as follows.

$$I_a = \begin{array}{c} \begin{array}{ccccccc} R_1 & R_2 & R_3 & R_5 & R_4 & R_7 & R_6 \end{array} \\ \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \end{array} \begin{array}{c} X_1 \\ X_2 \\ X_2 + X_3 \\ X_3 \\ 0 \\ 2X_3 \\ X_1 + X_3 \end{array}$$

Observe now that all rows now contain a -1 which make all complexes now reactant complexes. The corresponding network to the updated I_a is given below.



Despite making all complexes reactant complexes, the network remains to be not weakly reversible. An observation we can get from the I_a and its corresponding network is that the zero complex is not a product complex. If we add at least one of the following reactions below, the network becomes weakly reversible.

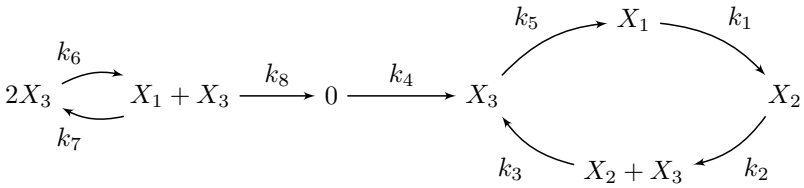
$$X_1 \longrightarrow 0, \qquad X_2 \longrightarrow 0, \qquad X_2 + X_3 \longrightarrow 0, \qquad X_3 \longrightarrow 0$$

This is equivalent to adding 1 in the 5th row of I_a. In such case, all complexes now are both a reactant complex and a product complex. So, the condition that all complexes of the network be a reactant and product

complex must be imposed for weak reversibility. However, we can see in the example below that it is not a sufficient condition for weak reversibility. Instead of adding at least one of the reactions above, consider adding either of the following reactions below (*Note: The complexes* $2X3$ *and* $X_1 + X_3$ *are from another linkage class.*).

$$2X_3 \longrightarrow 0, \qquad\qquad X_1 + X_3 \longrightarrow 0$$

Suppose we add the latter reaction. The network and its corresponding I_a become



$$I_a = \begin{array}{c} \begin{matrix} R_1 & R_2 & R_3 & R_5 & R_4 & R_8 & R_7 & R_6 \end{matrix} \\ \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 \end{bmatrix} \end{array} \begin{array}{c} \\ X_1 \\ X_2 \\ X_2 + X_3 \\ X_3 \\ 0 \\ 2X_3 \\ X_1 + X_3 \end{array}$$

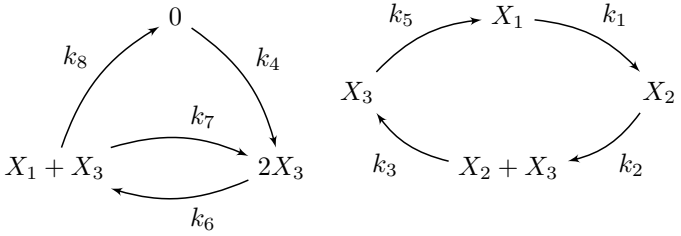
Here, all rows now contain entries of -1 and 1. The network, however, remains not to be weakly reversible. This implies that we should impose another condition to get the general structure of a weakly reversible I_a. Now, we can see that there exist two cycles in the network. The first one involves the complexes $X_1, X_2, X_2 + X_3$, and X_3 while the second one involves $2X_3$ and $X_1 + X_3$. Consider the fifth column. The complex X_3 serves as a product complex to a reaction whose reactant complex is a complex outside the cycle that involves X_3. Similarly, the complex $X_1 + X_3$ serves as a reactant to a reaction whose product complex is a

complex outside the cycle that involves $X_1 + X_3$. These reactions make the network not weakly reversible.

We can remedy this by removing reaction $0 \longrightarrow X_3$ and adding the reaction $0 \longrightarrow 2X3$. Doing so gives us an I_a as follows.

$$I_a = \begin{array}{c} \begin{matrix} R_1 & R_2 & R_3 & R_5 & R_4 & R_8 & R_7 & R_6 \end{matrix} \\ \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 \end{bmatrix} \end{array} \begin{matrix} X_1 \\ X_2 \\ X_2 + X_3 \\ X_3 \\ 0 \\ 2X_3 \\ X_1 + X_3 \end{matrix}$$

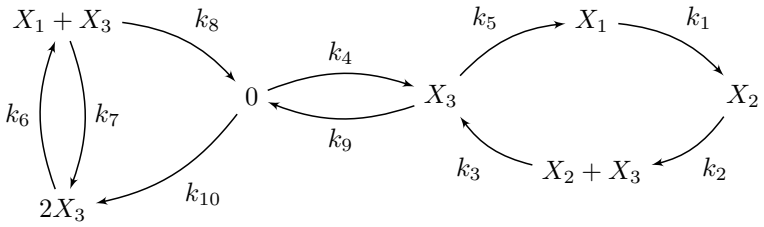
Notice that the I_a becomes a block matrix which corresponds to a network with two linkage classes. Here, all rows of each linkage class contain a -1 and 1. We can also see that no complex in any of the cycle is involved in a reaction with a complex outside the cycle. The reaction network corresponding to the I_a above is given below.



We then make the following remark with regards to the structure of a weakly reversible network: If we have a cycle, no complex in the cycle must be involved in a reaction with a complex outside the cycle. Equivalently, ***strongly connected components must correspond to a block matrix in I_a with zero off-diagonal blocks***.

Now, suppose we put the reaction $0 \longrightarrow X3$ back and add the reaction $X3 \longrightarrow 0$ while retaining the reaction $0 \longrightarrow 2X3$. In this way, the 0 complex becomes a part of the cycle and that the network now only contains

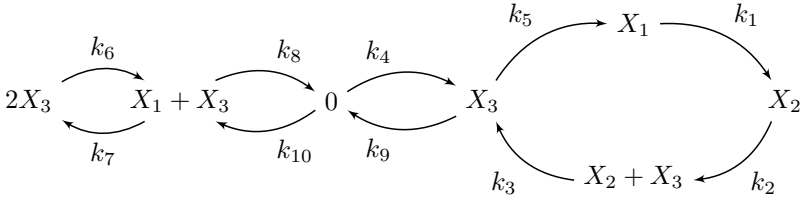
one linkage class. Refer to the updated network and its corresponding I_a below.



$$I_a = \begin{array}{c} \begin{matrix} R_1 & R_2 & R_3 & R_5 & R_4 & R_8 & R_7 & R_6 & R_9 & R_{10} \end{matrix} \\ \left[\begin{array}{cccccccccc} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 0 & 0 \end{array}\right] \end{array} \begin{matrix} X_1 \\ X_2 \\ X_2 + X_3 \\ X_3 \\ 0 \\ 2X_3 \\ X_1 + X_3 \end{matrix}$$

Here, the network still is weakly reversible. We make an additional remark with regards to the structure of a weakly reversible network: If we have a cycle, outside complexes attached to any complex in the cycle must both be a reactant and product to that cycle.

Now, we'll try to remove the reaction $0 \longrightarrow 2X3$. Instead, add the reaction $0 \longrightarrow X1 + X3$. The resulting network and its corresponding I_a are given below.



$$I_a = \begin{array}{c} \begin{matrix} R_1 & R_2 & R_3 & R_5 & R_4 & R_8 & R_7 & R_6 & R_9 & R_{10} \end{matrix} \\ \left[\begin{matrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 0 & 1 \end{matrix}\right] \end{array} \begin{matrix} X_1 \\ X_2 \\ X_2 + X_3 \\ X_3 \\ 0 \\ 2X_3 \\ X_1 + X_3 \end{matrix}$$

Notice that the sum of all entries of each row and column is 0. We can make the following remark: ***A row in*** I_a ***will have a sum of*** 0 ***if and only if there is an equal number of incoming and outgoing reactions for the corresponding complex.***

Now, it has been established that the condition that every row of I_a must have entries of -1 and 1 is needed for weak reversibility. Therefore, ***the dimension of an*** I_a ***that encodes a weakly reversible network must be*** $n \leq r$. Suppose $n > r$ (i.e., there are more complexes than reactions), we cannot have a -1 and 1 in every row. This is important for the algorithm because this serves as a terminality condition, as we will see in Section 4.2.

3.2 Strategy for narrowing down candidates for I_a

Our goal is to make the algorithm more efficient by narrowing down the possible candidates for I_a. To achieve this, we use the block matrix structure of I_a and an important result on positive-dependent networks. Note that the matrix I_a can be represented as a block matrix in the following form:

$$I_a = \begin{bmatrix} I_{a,1} & 0 & \dots & 0 \\ 0 & I_{a,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I_{a,\ell} \end{bmatrix},$$

where each $I_{a,i}$ corresponds to an individual linkage class, and the 0 entries represent zero matrices. The matrix structure for I_a will now be applied individually to each $I_{a,i}$, significantly reducing the possible candidates for I_a. On the other hand, Talabis and Mendoza [8] build on the result of Hong et al. [15] to establish the following theorem: if a network $\mathscr{N}$ is positive dependent, then, for any kinetics K, the system $(\mathscr{N}, K)$ has a weakly reversible network translation $(\mathscr{N}^{\#}, K^{\#})$. We will utilize this result to identify candidates for I_a by examining the network's positive dependency properties.

Our approach consists of the following steps:

1. Identify subnetworks within the overall network that are positive-dependent.
2. For each positive-dependent subnetwork, associate a corresponding matrix $I_{a,i}$. These matrices $I_{a,i}$ will represent the weakly reversible linkage classes of the network.
3. Construct I_a as a block matrix composed of the individual I_{a_i} blocks.

This strategy will allow us to systematically narrow down possible structures for I_a based on the network's positive dependencies. To identify the positive dependent subnetwork, we present the following definition and algorithm:

Definition 10. Let $\mathbf{R} = \{\mathbf{R}_1, \ldots, \mathbf{R}_r\}$ be the set of positive dependent reaction vectors. The positive dependent graph of $\mathbf{R}$ is the (undirected) graph $G = (V, E)$ with vertex set $V = \{v_1, \ldots, v_r\}$ and edge set E such that (v_i, v_j) is an edge in E if and only if there exist positive constants a_i, a_j and non-negative a_k's with $a_i\mathbf{R}_i + a_j\mathbf{R}_j + \sum_{k=1}^{r-2} a_k\mathbf{R}_k = 0$.

We now consider the following detailed method to obtain a positive dependent graph from a positive dependent network.

Remark 1. *We note that reaction vectors may not be unique, as two reactions may have the same reaction vectors. The following method is*

inspired by the work of Hernandez and De la Cruz [2], who employed a coordinate graph to find independent decompositions of chemical reaction networks. In this work, we adopt a similar strategy with variations to identify positive dependent decompositions.

Method of Finding a Positive Dependent Subgraph

1. Consider the set of reaction vectors $\{R_1, R_2 \ldots, R_r\}$.
2. Construct the vertex set of the positive dependent graph $G = (V, E)$ by representing each R_i as vertex v_i.
3. For each unconnected vertex v_i, consider the vector R_i. For $k = 1$ to $r - 1$:
   a. Check if there exist positive coefficients a_i's such that $a_i\mathbf{R}_i + \sum_{j=1}^{k} a_j\mathbf{R}_j = 0$.
   b. If such a combination is possible, add the edge (v_i, v_j) to E for each $\mathbf{R}_j$ in the combination. Go to Step 4.
4. Repeat Step 3 for each remaining unconnected vertex. Stop once all vertices are connected, yielding a positive dependent graph.

Remark 2. *Each connected subgraph in the positive-dependent graph corresponds to a positive-dependent subnetwork associated with a matrix $I_{a,i}$. Each $I_{a,i}$ matrix represents a weakly reversible linkage class(es) of the network.*

4 Constructing weakly reversible realizations

4.1 S-systems

Recall that the general rule of representing an S-system model as a reaction network according to [3] is as follows:

For each $\dot{X}_i$, we associate the influx reaction by $\sum X_{g,j} \to X_i + \sum X_{g,j}$ while we associate the efflux reaction by $X_i + \sum X_{h,j} \to \sum X_{h,j}$.

If we take the reaction vectors for these reactions, we get $-X_i$ and X_i. Therefore, the proposed algorithm for constructing a weakly reversible realization for S-system models is founded on these two rules:

Rule 1. *We associate the reaction $X_i + \sum_{k=1}^{m'} X_k \xrightarrow{\alpha_i} 2X_i + \sum_{k=1}^{m'} X_k$ for the influx terms where X_i's are the dependent species in the model, X_k's are the independent species, and m' is the number of independent species in the model.*

Rule 2. *We associate the reaction $2X_i + \sum_{k=1}^{m'} X_k \xrightarrow{\alpha_i} X_i + \sum_{k=1}^{m'} X_k$ for the efflux terms where X_i's are the dependent species in the model, X_k's are the independent species, and m' is the number of independent species in the model.*

The complete, step-by-step description of the algorithm for finding weakly reversible realization of S-systems is given by Algorithm 2.

Algorithm 2. *The S-system algorithm for constructing weakly reversible realization of BST models is as follows.*

1. *We start with the ODE form of the model. If we are given a biochemical map instead, we determine first its ODE form which is given by:*

$$\dot{X}_i = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}}$$

   for $i = 1, \ldots, n$ where n and m are the number of dependent and independent species, and g_{ij} and h_{ij} are the kinetic orders on the influx and efflux terms, respectively.

2. *For the influx terms of $\dot{X}_i$, perform **Rule** 1.*

3. *For the efflux terms of $\dot{X}_i$, perform **Rule** 2.*

4. *Stop if we have already done applying the method for all dependent species in the model.*

The two rules above ensures us that we get the same reaction vectors as the CRN representation according to the method in [3]. Consequently, we get the same stoichiometric matrix N. Equivalently, refer to Figure 2 for the flowchart version of Algorithm 2.

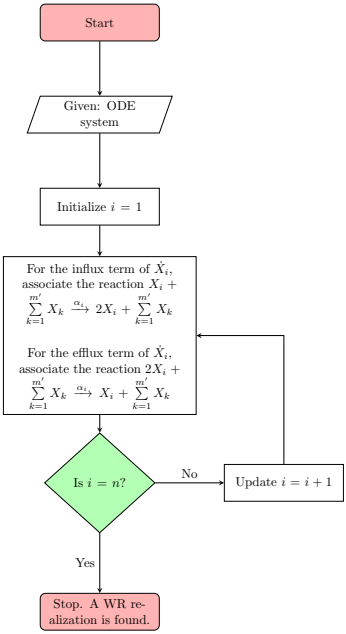


Figure 2. Flowchart of Algorithm 2

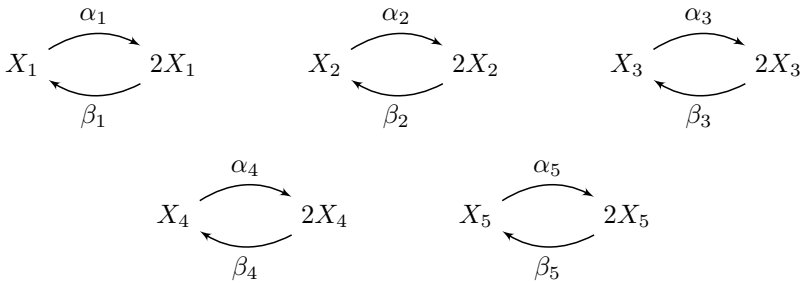
Remark 3. *Algorithm 2 guarantees that a weakly reversible realization can always be constructed for any S-system.*

Running Example 3 - Part 1. Consider the biochemical system model called HS96. It is a genetic network proposed by Hlavacek and Savageau in 1996 with 5 dependent variables and 10 reactions [18]. The ODE model

of the system under the S-system framework is given by

$$\begin{aligned}
\dot{X}_1 &= \alpha_1 X_3^{g_{13}} X_5^{g_{15}} - \beta_1 X_1^{h_{11}} \\
\dot{X}_2 &= \alpha_2 X_1^{g_{21}} - \beta_2 X_2^{h_{22}} \\
\dot{X}_3 &= \alpha_3 X_2^{g_{32}} - \beta_3 X_2^{h_{32}} X_3^{h_{33}} \\
\dot{X}_4 &= \alpha_4 X_3^{g_{43}} X_5^{g_{45}} - \beta_4 X_4^{h_{44}} \\
\dot{X}_5 &= \alpha_5 X_4^{g_{54}} - \beta_5 X_5^{h_{55}}
\end{aligned}$$

where g_{ij}'s and h_{ij}'s are the kinetic orders. Using Algorithm 2, the weakly reversible realization is given by:



where α_i's and β_i's are the reaction rate coefficients. The table below compares the network and kinetic properties of the new network (the weakly reversible realization) to that of the old network which is the representation based on [3].

Table 1. Comparison of network and kinetic properties (Running Example 3)

Network and Kinetic Properties	Algorithm 1	Algorithm 2
Species	5	5
Complexes	11	10
Reactant complexes	6	10
Reversible reactions	1	5
Irreversible reactions	8	0
Reactions	10	10
Linkage classes	2	5

Strong linkage classes	10	5
Terminal linkage classes	5	5
Rank	5	5
Reactant rank	5	5
Deficiency	4	0
Reactant deficiency	1	5
ILC	No	Yes
Weakly reversible	No	Yes
t-minimal	No	Yes
Positive dependence	Yes	Yes
Endotactic	No	Yes
PL-RDK	Yes	Yes
PL-TIK	Yes	Yes
Conservative	No	No
Concordance	No	No

The algorithm has a remarkable effect on the system: ILC, weak reversibility, t-minimality and edotacity are now attained. Also, observe that the deficiency of the newly constructed networks became 0. We formalize this result in the theorem below.

Theorem 8. *If we perform the S-systems algorithm (Algorithm 2) for constructing a weakly reversible realization, the resulting reaction network has* $\delta = 0$.

Proof. Recall that n is the number of complexes and l is the number of linkage classes in the reaction network.

We define m' to be the number of dependent species in the model. From the algorithm, it is clear that $n = 2m'$ and $l = m'$. Furthermore,

the stoichiometric matrix N is an $m' \times 2m'$ matrix of the form

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & \cdots & 0 & 0 \\ . & . & . & . & . & . & \cdots & . & . \\ . & . & . & . & . & . & \cdots & . & . \\ . & . & . & . & . & . & \cdots & . & . \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix}.$$

So, we can say that $\text{rank}(N) = m'$. Therefore, $\delta = n - l - s = 2m' - m' - m' = 0$. ■

We can now relate this to our previous results on PL-TIK systems (Theorems 1 and 3). We have the following:

Corollary 1. *Given an S-system with the induced kinetic system $(\mathscr{N}, K)$ by Algorithm 3, we have the following:*

i. $(\mathscr{N}, K)$ is PL-TIK if and only if $E_+(\mathscr{N}, K) \neq \emptyset$.

ii. If $(\mathscr{N}, K)$ is PL-TIK,

$$E_+(\mathscr{N}, K) = Z_+(\mathscr{N}, K) = \left\{ x \in \mathbb{R}^m_+ \,|\, \log(x) - \log(x^*) \in \left(\tilde{S}_R\right)^\perp \right\}$$

and $|E_+(\mathscr{N}, K) \cap \boldsymbol{Q}| = |Z_+(\mathscr{N}, K) \cap \boldsymbol{Q}| = 1$ for each positive kinetic class $\boldsymbol{Q}$.

Running Example 3 - Part 2. Considering the induced system of Algorithm 3, by definition, the system is PL-RDK. We define $\hat{T}$ (written as transpose) as follows

$$\hat{\mathbf{T}}^{\top} = \begin{bmatrix} 0 & 0 & 1.0041 & 0 & -1.0009 & 1 & 0 & 0 & 0 & 0 \\ 2.0113 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1.9919 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1.9963 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -0.9599 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -0.9702 & 1.8996 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1.9060 & 0 & -0.9288 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1.8127 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1.9749 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1.9556 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Kinetic orders are taken from the paper of Rinon et al. where they estimated the parameters of the S-sysem model using Hybrid Genetic Algorithm [18]. Here, $\text{rank}(\hat{T}) = 10$. Therefore, the system is PL-TIK (by Definition 3). Thus, by Corollary 1, we can say that there exists positive steady states for this system which are all complex balanced.

4.2 GMA systems

We will use the details discussed in Section 3 for the algorithm of GMA systems. For each possible I_a that encodes a weakly reversible network, we solve the induced linear system $Y \cdot I_a = N$ where the unknowns are the entries of Y.

In this formulation, the number of reactions r is fixed since we aim to get a similar N as the initial network. So, we start with the "simplest" I_a of weakly reversible networks which just contains the minimum number of complexes n for a given r until $n = r$. So, the complete, step-by-step description of the algorithm for finding weakly reversible realization of GMA systems is given below.

Algorithm 3. *The GMA system algorithm for constructing weakly reversible realization of BST models is as follows.*

1. *We start with the biochemical map of the model. If we are given an ODE form instead, we determine the biochemical map by considering*

the general structure of the GMA system model which is given by:

$$\dot{X}_i = V_i^+ - V_i^-$$

where V_i^+ *and* V_i^- *are the sums of all incoming and outgoing processes (reactions) of* X_i*, respectively.*

2. *Construct the CRN representation of the system based on the method in [3].*

3. *Break down network into positive-dependent subnetworks (Refer to Section 3.2). Determine the stoichiometric matrix of each subnetwork.*

4. *For each subnetwork, construct an incidence matrix* $I_{a_{n\times r}}$ *that encodes a weakly reversible network (Refer to Section 3.1) such that* n *is the minimum number of complexes for a fixed* r*.*

$$\begin{gathered}
\textit{Number of reactions} \rightarrow \textit{Minimum number of complexes} \\
2 \rightarrow 2 \\
(2, 6] \rightarrow 3 \\
(6, 12] \rightarrow 4 \\
(12, 20] \rightarrow 5 \\
\vdots \\
((n-1)\cdot(n-2), n\cdot(n-1)] \rightarrow n
\end{gathered}$$

5. *Given the* N *from step 3 and* I_a *from step 4, solve the matrix equation*

$$Y \cdot I_a = N$$

If a solution exists, Y *and* I_a *is the molecularity and incidence matrices of the constructed weakly reversible realization, respectively. Repeat for all subnetworks. Otherwise, go back to step 4. Stop if* $n = r$*. If no solution exists for all* I_a*, the algorithm cannot find a weakly reversible realization of the CRN.*

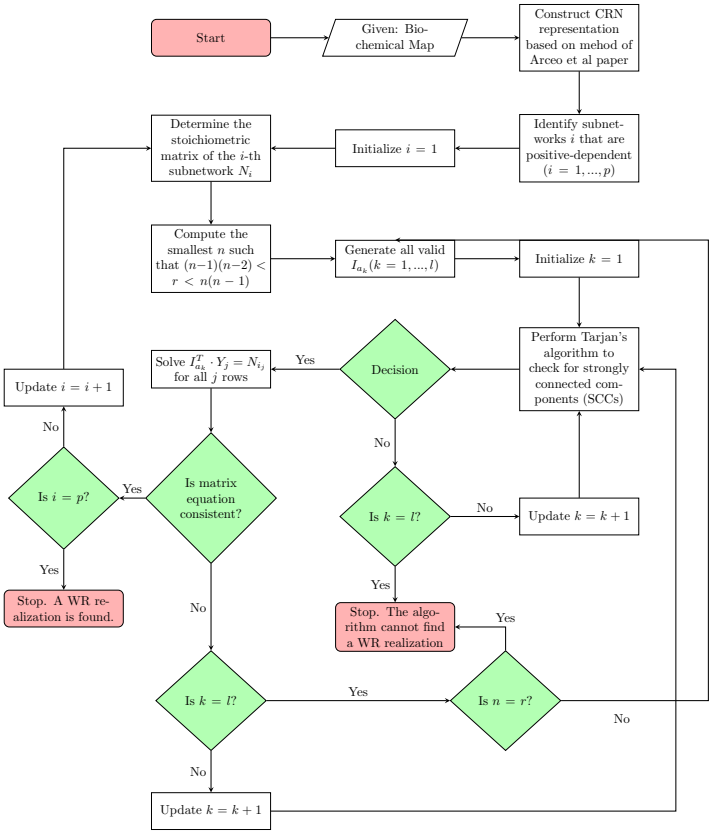


Figure 3. Flowchart of Algorithm 3

For the complete details and step-by-step procedure of Algorithm 3, refer to Figure 3. In generating valid I_a, the requirements are: 1) Each column must only have two nonzero entries: -1 and 1, 2) Each row must have at least one -1 entry and one 1 entry, and 3) No two columns must be the same.

Furthermore, in the decision after performing Tarjan's algorithm to check for SCCs, both must be satisfied: 1) Do all SCCs contain more than one complex? and 2) Do the SCCs correspond to block submatrices in I_a with zero off-diagonal blocks? Note that if there is only one SCC, condition 2 is vacuously true.

Remark 4. *It is worth noting that the resulting weakly reversible network after performing Algorithm 3, if successful, can either be PL-NDK or PL-RDK. Therefore, one can iteratively construct I_a (as outlined in Step 4) until the desired kinetic properties, such as PL-RDK, are achieved. This is one advantage of Algorithm 3 over the method in [15].*

Running Example 4 - Part 1. Consider a linear pathway with end-product inhibition (refer to [3]). According to the method in [3], the chemical reaction network corresponding to the system under the GMA framework is given by

$$\begin{aligned} 0 &\longrightarrow X_4 & X_3 &\longrightarrow 0 \\ X_1 &\longrightarrow X_2 & X_2 &\longrightarrow X_3 \\ X_4 + X_3 &\longrightarrow X_1 + X_3 \end{aligned}$$

We cannot find a positive dependent subgraph for this particular subnetwork. So, we take the whole network as is. We then detemine the stoichiometric matrix which is given by

$$N = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & -1 & 0 & 0 & 0 \end{bmatrix}.$$

Since $r = 5$, the minimum number of complexes is 3. However, there is no valid incidence matrix that encodes a weakly reversible network that produces a consistent matrix equation $Y \cdot I_a = N$ with 3 and 4 complexes. So, we consider an incidence matrix with 5 complexes.

$$I_a = \begin{bmatrix} -1 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}.$$

This induces the matrix equation

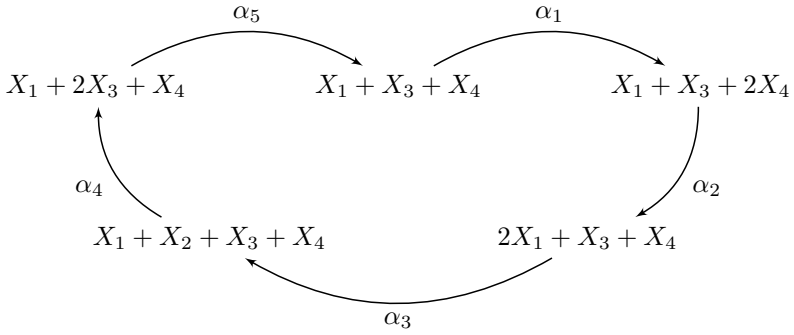
$$\begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \\ a_5 & b_5 & c_5 & d_5 \end{bmatrix}^\top \begin{bmatrix} -1 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix}^\top$$

which yields the following molecularity matrix Y if we set $a_5 = 1, b_5 = 0, c_5 = 2$, and $d_5 = 1$ given below

$$Y = \begin{array}{c} \begin{matrix} C_1 & C_2 & C_3 & C_4 & C_5 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 2 \\ 1 & 2 & 1 & 1 & 1 \end{bmatrix} \end{array} \begin{matrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{matrix} .$$

Instead of considering a single large coefficient matrix A, solving the matrix equation above can be done by working on one row at a time to save computing time. This leads us to the following remark: ***The rows of Y are independent from each other***. In other words, stoichiometric coefficients of species X_i in complex C_j can be determined independently. Furthermore, ***the existence of free variables implies that the weakly reversible realization for GMA system models may not be unique***.

Using the molecularity matrix Y above, the resulting network is given by



where the α_i's are the reaction rate coefficients. The network now becomes weakly reversible.

Table 2. Comparison of network and kinetic properties (Running Example 4)

Network and Kinetic Properties	Algorithm 1	Algorithm 3
Species	4	4
Complexes	7	5
Reactant complexes	5	5
Reversible reactions	0	0
Irreversible reactions	5	5
Reactions	5	5
Linkage classes	2	1
Strong linkage classes	7	1
Terminal linkage classes	2	1
Rank	4	4
Reactant rank	4	4
Deficiency	1	0
Reactant deficiency	1	1
ILC	No	Yes
Weakly reversible	No	Yes
t-minimal	Yes	Yes
Positive dependence	Yes	Yes
Endotactic	No	Yes
PL-RDK	Yes	Yes
PL-TIK	Yes	Yes
Conservative	No	No
Concordance	Yes	No

Building on Theorems 1 and 3, we have the following corollary:

Corollary 2. *Given a GMA-system with the induced kinetic system* $(\mathscr{N}, K)$ *by Algorithm 3, we have the following:*

i. $(\mathscr{N}, K)$ *is PL-TIK if and only if* $Z_+(\mathscr{N}, K) \neq \emptyset$.

ii. If $(\mathscr{N}, K)$ *is PL-TIK,*

$$Z_+(\mathscr{N}, K) = \left\{ x \in \mathbb{R}^m_+ | \log(x) - \log(x^*) \in \left(\tilde{S}_R\right)^\perp \right\}$$

and $|Z_+(\mathscr{N}, K) \cap \boldsymbol{Q}| = 1$ *for each positive kinetic class* $\boldsymbol{Q}$*.*

iii. If $(\mathscr{N}, K)$ *is PL-TIK and* $\delta = 0$*,* $E_+(\mathscr{N}, K) = Z_+(\mathscr{N}, K)$*.*

Running Example 4 - Part 2. From our Running Example 4, direct computations show that the tranformed system is PL-RDK. Determining whether it is PL-TIK, we define $\hat{T}$ by

$$\hat{T} = \begin{bmatrix} 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 0.75 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

The system is PL-TIK since $\text{rank}(\hat{T}) = 5$ (by Definition 3). Furthermore, given that $\delta = 0$, from Corollary 2, we can say that there exists positive steady states which are complex balanced.

Algorithm 3 becomes impractical for large systems since it entails an enumeration of incidence matrices that encode weakly reversible network especially if we cannot find sufficiently small positive-dependent subgraphs for the whole network.

4.3 Large GMA systems

One remedy we can do for the impracticality of using Algorithm 3 in constructing a weakly reversible realization for large GMA systems is to consider another type of kinetics. In particular, we consider PYK for constructing a weakly reversible realization of large GMA systems. Reaction networks endowed with PYK are easier to work with since we can think

of the ODE model of the network as follows.

$$\dot{X}_i = \left(\sum_{i=1}^{r^+} \prod_{j\in(\text{influx})_i} X_j\right) - \left(\sum_{k=1}^{r^-} \prod_{g\in(\text{efflux})_k} X_g\right)$$

where r^+ is the number of incoming reactions to X_i, r^- is the number of outgoing reactions from X_i, X_j are the contributing species to the influx terms, and X_g are the contributing species to the efflux terms.

In this way, we can think of it as only consisting of "two" terms such that the first one is for influx terms (i.e., the sum of the product of contributing species for each incoming reaction) and the second one is for the efflux terms (i.e., the sum of the product of contributing species for each outgoing reaction). Therefore, we can apply Algorithm 2 but with PYK.

Since Algorithm 3 becomes impractical for large GMA systems, we can instead model the system using PYK and apply the method for constructing weakly reversible realizations for S-systems. The complete, step-by-step description of the algorithm for finding a weakly reversible realization of large GMA systems is given below.

Algorithm 4. *The algorithm for large GMA systems in constructing a weakly reversible realization (with Poly-PL kinetics) of BST models is as follows.*

1. *We start with the ODE form of the model (must be of the form Equation 4.3). If we are given a biochemical map instead, we determine first its ODE form which is given by:*

$$\dot{X}_i = \sum_{j\in influx} v_{ji} - \sum_{k\in efflux} v_{ik}$$

where $v_{ji} = \prod_{j\in influx} X_j$ *and* $v_{ik} = \prod_{k\in efflux} X_k$ *are the corresponding power law representations of each of the influx and efflux reactions, respectively.*

2. *For the influx terms of* $\dot{X}_i$*, perform* ***Rule*** *1.*

3. *For the efflux terms of* $\dot{X}_i$*, perform* ***Rule*** *2.*

4. *Stop if we have already done applying the method for all dependent species in the model.*

Note that Algorithm 4 is essentially Algorithm 2 but with poly-PL kinetics. Therefore, the resulting network after applying Algorithm 4 would have deficiency 0. We then state the following theorem:

Theorem 9. *If we perform the algorithm for large GMA systems modelled with Poly-PL kinetics (Algorithm 4) for constructing a weakly reversible realization, the resulting reaction network has* $\delta = 0$.

The proof is similar to that of Theorem 8. Furthermore, from the PY-TIK results, we have the following:

Corollary 3. *Given a GMA-system with the induced poly-PL system* $(\mathscr{N}, K)$ *by Algorithm 4, we have the following:*

i. $(\mathscr{N}, K)$ *is PY-TIK if and only if* $Z_+(\mathscr{N}, K) \neq \emptyset$.

ii. If $(\mathscr{N}, K)$ *is PY-TIK,*

$$Z_+(\mathscr{N}, K) = \left\{ x \in \mathbb{R}^m_+ | \log(x) - \log(x^*) \in \left(\tilde{S}_j\right)^\perp \right\}$$

and $|Z_+(\mathscr{N}, K) \cap \boldsymbol{Q}| = 1$ *for each positive kinetic reactant flux class* $\boldsymbol{Q}_j$.

5 Applications: Concentration robustness and carbon cycle models

In this section, we present several applications of the algorithms discussed previously. First, we will discuss how the proposed algorithms help to identify key features of the system, such as concentration robustness. We will then conclude with an application on a GMA representation of the carbon cycle models by Anderies et al [17] and Heck et al [32].

5.1 Concentration robustness in weakly reversible systems

The concept of absolute concentration robustness (ACR) was first introduced by Shinar and Feinberg [13]. ACR pertains to a phenomenon in which a species in a chemical kinetic system carries the same value for any positive steady state the network may admit regardless of initial conditions. In particular, a kinetic system $(\mathscr{N}, K)$ has ACR in a species $X \in \mathscr{S}$ if there exists $c^* \in E_+(\mathscr{N}, K)$ and for every other $c^{**} \in E_+(\mathscr{N}, K)$, we have $c^{**}_X = c^*_X$. Fortun and Mendoza [24] introduced the concept of balanced concentration robustness (BCR). A complex balanced chemical kinetic system $(\mathscr{N}, K)$ has **balanced concentration robustness (BCR)** in a species $X \in \mathscr{S}$ if X has the same value for all $c \in Z_+(\mathscr{N}, K)$. This is another type of concentration robustness that is weaker than ACR. For the PLK systems (or subsystems), the key property for ACR (and BCR) in a species X is the presence of an SF-reaction pair, which is defined as follows.

Definition 11. A pair of reactions in a PLK system is called a **Shinar-Feinberg pair** (or **SF-pair)** in a species X if their kinetic order vectors differ only in X. A subnetwork of the PLK system is of **SF-type** if it contains an SF-pair in X.

The following are key results on ACR and BCR from Fortun and Mendoza [24].

Theorem 10 (Theorem 6 of [24])**.** *Let $(\mathscr{N}, K)$ be a deficiency zero PL-RDK with a positive equilibrium. If a pair of reactions in a linkage class forms an SF-pair species X, then it has ACR in X.*

Proposition 11 (Proposition 7 of [24])**.** *Let $(\mathscr{N}, K)$ be a kinetic system with a positive equilibrium and an independent decomposition $\mathscr{N} = \mathscr{N}_1 \cup \mathscr{N}_2 \cup \cdots \cup \mathscr{N}_p$. If there is an $\mathscr{N}_i$ with PL-RDK $(\mathscr{N}_i, \mathscr{K}_i)$ of SF-type in $X \in \mathscr{S}$ such that $\delta \leq 1$ then $(\mathscr{N}, K)$ has ACR in X.*

Proposition 12 (Proposition 8 of [24])**.** *Let $(\mathscr{N}, K)$ be a kinetic system with a complex balanced equilibrium and an incidence independent decom-*

position $\mathscr{N} = \mathscr{N}_1 \cup \mathscr{N}_2 \cup \cdots \cup \mathscr{N}_p$. If there is an $\mathscr{N}_i$ with PL-RDK $(\mathscr{N}_i, \mathscr{K}_i)$ of SF-type in $X \in \mathscr{S}$ such that $\delta \leq 1$ then $(\mathscr{N}, K)$ has BCR in X.

In previous works, deficiency has been a necessary condition for the ACR/BCR results. The following result allows us to eliminate this condition. However, before presenting the theorem, we define the following

Definition 12. An SF-pair $\{r, r'\}$ in X is called consecutive if the product complex of r is the reactant complex of r'.

Theorem 13. *Let $(\mathscr{N}, K)$ be a weakly reversible PL-TIK system. If the system has a consecutive SF-pair in X, then it has balanced concentration robustness in X.*

Proof. By Theorem 6 of [7], $Z_+(\mathscr{N}, K) \neq \emptyset$ for all rate constants and

$$Z_+(\mathscr{N}, K) = \left\{ x \in \mathbb{R}^m_{\geq} \middle| \log(x) - \log(x^*) \in (\widetilde{S}_R)^{\perp} \right\}.$$

Let $c^*, c^{**} \in Z_+(\mathscr{N}, K)$. Hence, $\log(c^*) - \log(c^{**}) \in (\widetilde{S}_R)^{\perp}$. Note that $\widetilde{S}_R$ is the space generated by $T(I_a)$ and each column of $T(I_a)$ corresponds to a reaction (i, j) in such a way that the corresponding column is $T_{\cdot,j} - T_{\cdot,i}$. Thus, we have

$$\langle T_{\cdot,j} - T_{\cdot,i}, \log(c^*) - \log(c^{**}) \rangle = 0, \quad \forall (i,j) \in \mathscr{R}. \tag{1}$$

Since $(i, j), (j, k) \in \mathscr{R}$ is an **SF-pair** in $X \in \mathscr{S}$, $T_{\cdot,j} - T_{\cdot,i} = aX \quad \forall a \in \mathbb{R}_+$.

Thus, Equation 1 will be reduced to $a(\log(c^*_X) - \log(c^{**}_X)) = 0$. Hence, $c^*_X = c^{**}_X$. That is, the system has balanced concentration robustness on X. ■

Remark 5. *Since Algorithms 2 and 4 generate zero deficiency networks, Theorem 10 and Propositions 11 and 12 can be used to conclude ACR and BCR. On the other hand, Theorem 13 can be applied for Algorithm 3 for BCR conclusion.*

5.2 Weakly reversible translates of the Anderies et al. model

The pre-industrial carbon cycle model by Anderies et al. [17] is a simple mass balance model involving three interacting carbon pools: land, atmosphere, and ocean. This system can be visually represented using a biochemical map, where nodes represent carbon pools, solid arrows indicate carbon transfer, and dashed arrows indicate modulation effects by the pools. Figure 4 presents the biochemical map of this model.

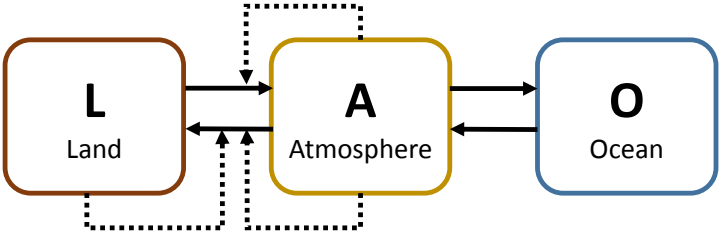


Figure 4. Biochemical map of Anderies et al.'s carbon cycle model in the pre-industrial state

In [23], Fortun et al. reviewed the model's design and underlying assumptions, detailing the parameters and ordinary differential equations in the pre-industrial state of the carbon cycle model. They approximated all rate processes using products of power law functions to derive a Generalized Mass Action (GMA) system approximation of the original model. The resulting ODEs from the approximation are given by:

$$\begin{cases} \dot{L} &= k_1 L^{p_1} A^{q_1} - k_2 L^{p_2} A^{q_2} \\ \dot{A} &= k_2 L^{p_2} A^{q_2} - k_1 L^{p_1} A^{q_1} - a_m A + a_m \beta O \\ \dot{O} &= a_m A - a_m \beta O, \end{cases} \tag{2}$$

For the case where the human terrestrial carbon off-take term (which accounts for human activities that reduce the capacity of the terrestrial pool to capture carbon, such as deforestation and land-use change) vanishes, the estimated kinetic orders are $p_1 = p_2 = -69$, $q_1 = 0.580148$ and

$q_2 = 0.910864$. Using the BST algorithm, the CRN representation for the model is:

$$\begin{aligned} L + 2A &\rightarrow 2L + A \\ L + A &\rightarrow 2A \\ A &\rightleftarrows O \end{aligned} \tag{3}$$

Its associated kinetic order matrix is:

$$F = \begin{array}{c} \\ r_1 \\ r_2 \\ r_3 \\ r_4 \end{array} \begin{array}{c} \begin{array}{ccc} L & A & O \end{array} \\ \begin{bmatrix} -69 & 0.580148 & 0 \\ -69 & 0.910864 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}. \tag{4}$$

Note that Theorems 10, 13 and Propositions 11, 12 cannot be used for ACR and BCR analysis because the network is not weakly reversible. Therefore, we employ the following algorithms.

5.2.1 Application of Algorithm 3

From the CRN representation of the carbon cycle model in 3, we determine its stoichiometric matrix which is given by

$$N = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

Since there are only 4 reactions which is manageable, we can bypass the step in finding positive-dependent subgraphs. We start identifying the incidence matrices I_a that encode weakly reversible networks of size 3×4. Only one valid I_a produces a consistent system which is given by the following matrix

$$I_a = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

Solving the induced linear system produces a realization that is PL-NDK. To apply the theory developed by Talabis et al. in [5, 7] and the parametrization results of Muller and Regensburger, we proceed with the algorithm to find a realization that is PL-RDK. Considering the incidence matrix given by

$$I_a = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

yields the matrix equation

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

Solving this linear system gives us six free variables: a_2, a_4, b_2, b_4, c_2, and c_4. Setting $a_2 = a_4 = b_2 = 1$, $c_2 = c_4 = 2$, and $b_4 = 0$, we have the following weakly reversible network:

$$2A + 2O \underset{k_2}{\overset{k_1}{\rightleftarrows}} L + A + 2O \qquad L + A + O \underset{a_m\beta}{\overset{a_m}{\rightleftarrows}} L + 2O$$

with kinetics $K(x)$ defined as

$$K(x) = \begin{bmatrix} k_1 L^{p_1} A^{q_1} \\ k_2 L^{p_2} A^{q_2} \\ a_m A \\ a_m \beta O \end{bmatrix}.$$

Furthermore, calculating the deficiency of the new network gives us $\delta = n - l - s = 4 - 2 - 2 = 0$. Therefore, we got a weakly reversible deficiency zero network after applying Algorithm 3 which is also PL-RDK.

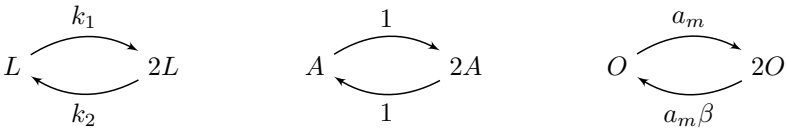
Consider the reactions $2A + 2O \leftrightarrow L + A + 2O$. We can say that this pair of reactions is an SF-pair in A since their kinetic orders differ only in A based on Equation 4. Furthermore, this SF-pair is consecutive by Definition 12. Therefore, there is a BCR in A by Theorem 13. But since the constructed network has deficiency zero, we can conclude that it has ACR in A.

Finally, we now apply the parametrization results of Muller of Regensburger. As mentioned, we can do this because our realization is weakly reversible and deficiency zero that is also a PL-RDK system. So, the monomial parametrization is given by $\mathbf{X} = \begin{bmatrix} k_1^{-r} k_2^r \xi & \frac{k_1^{-r} k_2^r}{\beta} & 1 \end{bmatrix}^\top$

where $r = \frac{1}{0.580148 - 0.910864}$ and $\xi \in \mathbb{R}_>$. The parametrization verifies the conclusion that there is an ACR in species A. In fact, there is also an ACR in species O.

5.2.2 Application of Algorithm 4

Using Algorithm 4, the weakly reversible realization is given by:



The kinetics $K(x)$ is defined as follows:

$$K(x) = \begin{bmatrix} k_1 L^{p_1} A^{q_1} \\ k_2 L^{p_2} A^{q_2} \\ k_2 L^{p_2} A^{q_2} + a_m \beta O \\ k_1 L^{p_1} A^{q_1} + a_m A \\ a_m A \\ a_m \beta O \end{bmatrix}.$$

The partition $P = \{\{L \leftrightarrow 2L\}, \{A \leftrightarrow 2A\}, \{O \leftrightarrow 2A\}\}$ induces a weakly reversible decomposition under poly-PL kinetics. Consider the 1st partition (under PL-RDK). The reactions $L \leftrightarrow 2L$ have the kinetic orders

$$\begin{matrix} r_1 \\ r_2 \end{matrix} \begin{bmatrix} -69 & 0.580148 & 0 \\ -69 & 0.910864 & 0 \end{bmatrix},$$

and hence, they form an SF-pair. Since the network deficiency is zero, by Theorem 10, we have an ACR in species A for the subnetwork $\{L \leftrightarrow 2L\}$. This result extends to the entire system, as P also induces an independent decomposition.

5.3 Weakly reversible translate of the Heck et al. model

The global carbon cycle model of Heck et al. [32] is built from the model of Anderies et al. [17]. In their revised model, the atmosphere-land interactions are modified for a better representation of empirically observed Earth system carbon dynamics [32]. Furthermore, they extended the model of Anderies et al. [17] by incorporating a societal intervention process called terrestrial carbon dioxide removal (tCDR) which mimics current international policies on climate change. This intervention sequesters and permanently stores terrestrial carbon in a carbon engineering sink. The model also considers pooling the geological carbon pool and the new sink to form a passive carbon pool and decoupling the atmospheric carbons into two nodes [25]. The biochemical map of the system is shown in Figure 5. A detailed discussion on the model development and calibration can be found in [32].

Using the BST algorithm [3] for GMA systems, the CRN representation of the model is:

$$\begin{aligned}
&A_1 + 2A_2 \to 2A_1 + A_2 && A_1 + 2A_4 \to 2A_1 + A_4 \\
&A_1 + A_2 \to 2A_2 && A_1 + A_4 \to 2A_4 \\
&A_2 \to A_3 && A_4 \to A_3 \\
&A_3 \to A_2 && A_3 \to A_4 \\
&A_4 + A_5 \to 2A_4 && A_1 + A_2 + A_4 \to A_5 + A_2 + A_4.
\end{aligned}$$

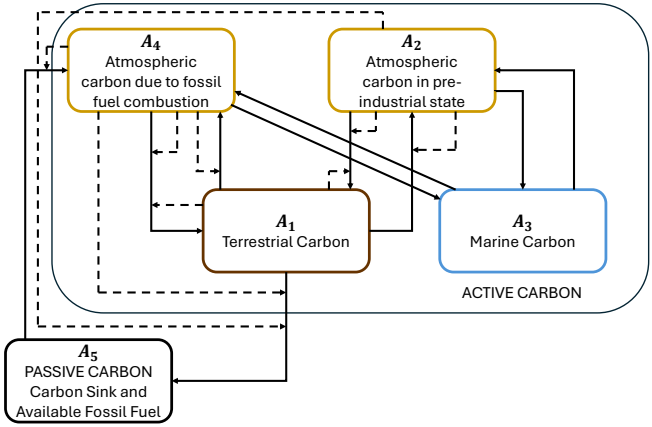


Figure 5. Biochemical map of Heck et al.'s carbon cycle model [25].

Its associated kinetic order matrix is given by

$$F = \begin{array}{c} \\ R_1 \\ R_2 \\ R_3 \\ R_4 \\ R_5 \\ R_6 \\ R_7 \\ R_8 \\ R_9 \\ R_{10} \end{array} \begin{array}{c} \begin{array}{ccccc} A_1 & A_2 & A_3 & A_4 & A_5 \end{array} \\ \begin{bmatrix} 199.75 & -86.03 & 0 & 0 & 0 \\ 159.84 & -63.32 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1.54 \\ -43.80 & 0 & 0 & 21.42 & 0 \\ -56.13 & 0 & 0 & 22.19 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 4.44 & 0 & 11.52 & 0 \end{bmatrix} \end{array}.$$

Note that the network above is not weakly reversible. To use the results for ACR and BCR analysis, we find a weakly reversible realization.

5.3.1 Application of Algorithm 3

From the CRN representation above, we have the following reaction vectors:

$$\begin{aligned} R_1 &: A_1 - A_2 & R_6 &: A_1 - A_4 \\ R_2 &: A_2 - A_1 & R_7 &: A_4 - A_1 \\ R_3 &: A_3 - A_2 & R_8 &: A_3 - A_4 \\ R_4 &: A_2 - A_3 & R_9 &: A_4 - A_3 \\ R_5 &: A_4 - A_5 & R_{10} &: A_5 - A_1 \end{aligned}$$

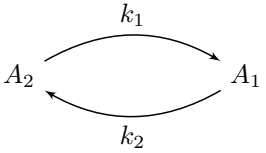
Here, we apply the positive dependent subgraph method discussed in Section 3.2. Note that $R_1 + R_2 = 0$, $R_3 + R_4 = 0$, and $R_5 + 2R_6 + R_7 + R_8 + R_9 + R_{10} = 0$. We can take $\mathscr{R}_1 = \{R_1, R_2\}$, $\mathscr{R}_2 = \{R_3, R_4\}$, and $\mathscr{R}_3 = \{R_5, R_6, R_7, R_8, R_9, R_{10}\}$ as the positive-dependent subgraphs of the whole network where $\mathscr{R}_1, \mathscr{R}_2$, and $\mathscr{R}_3$ are the reaction sets of the subgraphs. We first consider $\mathscr{R}_1$ whose induced subnetwork is given by

$$\begin{aligned} A_1 + 2A_2 &\to 2A_1 + A_2 \\ A_1 + A_2 &\to 2A_2. \end{aligned}$$

Considering the incidence matrix $I_a = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$, we solve

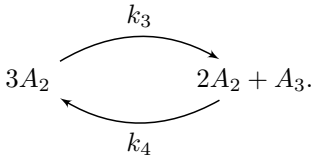
$$\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

whose solution gives us two free variables which are a_2 and b_2. Setting $a_2 = 1$ and $b_2 = 0$, we have the following weakly reversible subnetwork:



Similarly, we do this for $\mathscr{R}_2$ by considering the same incidence matrix.

Solving the induced matrix equation and setting $a_2 = 2$ and $b_2 = 1$, we get



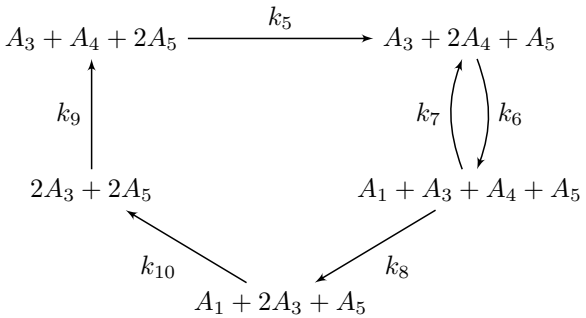
Lastly, we consider $\mathscr{R}_3$ whose induced subnetwork is given by

$$\begin{aligned} &A_4 + A_5 \to 2A_4 && A_4 \to A_3 \\ &A_1 + 2A_4 \to 2A_1 + A_4 && A_3 \to A_4 \\ &A_1 + A_4 \to 2A_4 && A_1 + A_2 + A_4 \to A_5 + A_2 + A_4 \end{aligned}$$

There are 6 reactions, however no valid incidence matrix that encodes a weakly reversible network produces a consistent matrix equation $Y \cdot I_a = N$ with 3 and 4 complexes. So, we consider

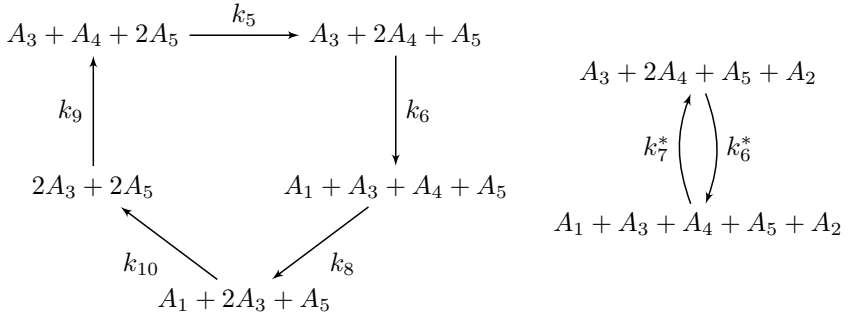
$$I_a = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 1 \\ 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

which gives us the weakly reversible subnetwork below by setting $a_5 = 1$, $b_5 = 2$, $c_5 = 1$, and $d_5 = 0$.

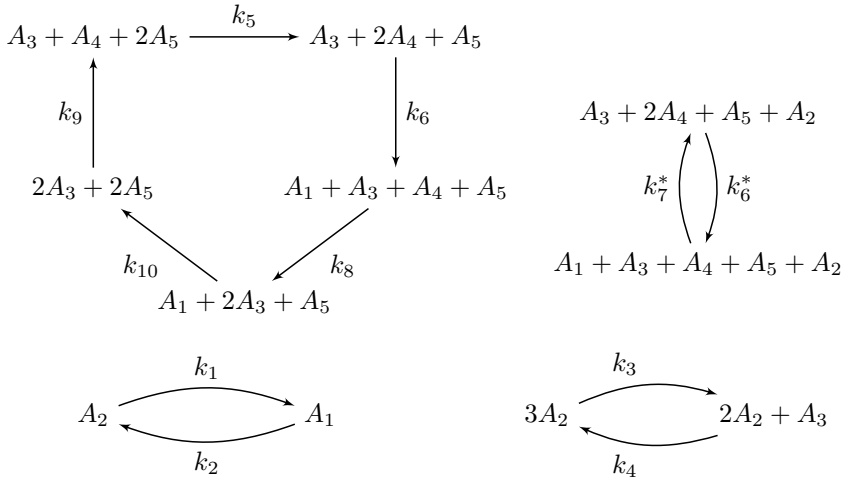


Notice that this particular subnetwork is PL-NDK due to the branching of reactions at the complex $A_1+A_3+A_4+A_5$. Using the CF-WR algorithm

introduced by Talabis and Mendoza [8], the kinetic system is now PL-RDK. We have the following:



With the obtained subnetworks above, the weakly reversible realization of the CRN representation of the Heck et al model [32] is given by



Here, direct computation shows that the linkage class decomposition can be considered as incidence independent decomposition. (See Definition 9.) Since each subnetwork is weakly reversible and of PL-TIK type, by Theorem 3, $Z_+(\mathscr{N}_i, K) \neq 0$ for all subnetworks. Consequently, by Theorem 7, $Z_+(\mathscr{N}, K) \neq 0$.

Now, consider R_5 (with rate constant k_5) and R_8 (with rate constant k_8). Note that both are part of a single linkage class. Referring to the kinetic order matrix, their kinetic orders differ only in species A_5. Hence,

R_5 and R_8 is an SF-pair in A_5. The deficiency of this particular subnetwork is 1. Therefore, by Proposition 12, the Heck system has BCR in species A_5.

6 Summary

Weak reversibility is an important property that reaction networks must possess because many results in CRNT requires this condition such as on the non-emptiness of the system's sets of positive and complex balanced equilibria, parametrization of these steady states, among others. However, BST models are not known to be weakly reversible. Hence, this study aims to construct weakly reversible realizations of BST models and apply known results in CRNT. To summarize, here are our main results:

1. This study considers two variants of BST models which are the GMA system and S-system. Building upon the method proposed by Arceo et al. [3] to represent these models as a chemical reaction network with power law kinetics, this paper develops algorithms to construct their weakly reversible realization.

2. For S-systems, we develop an algorithm that represents the influx and efflux terms of the ODE model as reversible reactions, ensuring that the network is weakly reversible. We also establish that the resulting deficiency is zero. On the other hand, the method for GMA systems is based on determining the entries of the molecularity matrix Y by solving the induced linear system from the matrix equation $Y \cdot I_a = N$. I_a is chosen in such a way that the associated network is weakly reversible.

3. Narrowing down the possible candidates for I_a is necessary to make the GMA algorithm more efficient. To achieve this, we build on the results of Talabis and Mendoza [8] and Hong et al. [15] which states that if a network that is positive-dependent then it has a weakly reversible realization. The approach then is to break down the entire network into positive-dependent subgraphs and focus on each subnetworks.

4. Another approach to address the impracticality of the GMA algorithm to large BST models is by considering Poly-PL kinetics and applying the S-system algorithm, thereby expanding the utility of the Poly-PL kinetics introduced in [6].

5. Because the resulting networks after applying the algorithms are weakly reversible and deficiency zero, we can apply the power law results of Talabis et al. [5–7] and parametrization results of Müller and Regensburger [30, 31].

6. Lastly, we apply the algorithms proposed in this paper for concentration robustness analysis. We demonstrate the results by considering the carbon cycle models by Anderies et al. [17] and Heck et al. [32].

References

[1] B. O. Palsson *Systems Biology: Constraint–Based Reconstruction and Analysis*, Cambridge Univ. Press, Cambridge, 2015.

[2] B. S. Hernandez, R. J. L. De la Cruz, Independent decompositions of chemical reaction networks, *Bull. Math. Biol.* **83** (2021) #76.

[3] C. P. Arceo, E. C. Jose, A. Marin–Sanguino, E. R. Mendoza, Chemical reaction network approaches to biochemical systems theory, *Math. Biosci.* **269** (2015) 135–152.

[4] C. P. Arceo, E. C. Jose, A. R. Lao, E. R. Mendoza, Reaction networks and kinetics of biochemical systems, *Math. Biosci.* **283** (2017) 13–29.

[5] D. A. S. Talabis, C. P. P. Arceo, E. R. Mendoza, Positive equilibria of a class of power law kinetics, *J. Math. Chem.* **56** (2018) 358–394.

[6] D. A. S. Talabis, D. M. Magpantay, E. R. Mendoza, E. G. Nocon, E. C. Jose, Complex balanced equilibria of weakly reversible poly-PL kinetic systems and evolutionary games, *MATCH Commun. Math. Comput. Chem.* **80** (2020) 375–402.

[7] D. A. S. Talabis, E. R. Mendoza, E. C. Jose, Complex balanced equilibria of weakly reversible power law kinetic systems, *MATCH Commun. Math. Comput. Chem.* **82** (2019) 601–624.

[8] D. A. S. Talabis, E. R. Mendoza, Network transformation-based analysis of biochemical systems, *J. Math. Biol.* **89** (2024) #62.

[9] E. C. Jose, D. A. S. Talabis, E. R. Mendoza, Absolutely complex balanced kinetic systems, *MATCH Commun. Math. Comput. Chem.* **88** (2022) 397–436.

[10] E. O. Voit, Biochemical systems theory: a review, *ISRN Biomath.* **6** (2013) 1–53.

[11] E. O. Voit, H. A. Martens, S. W. Omholt, 150 years of mass action law, *PLOS Comput. Biol.* **11** (2015) 1–7.

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

[13] G. Shinar, M. Feinberg, Structural sources of robustness in biochemical reaction networks, *Science* **327** (2010) 1389–1391.

[14] G. Szederkenyi, K. M. Hangos, Z. Tuza, Finding weakly reversible realizations of chemical reaction networks using optimization, *MATCH Commun. Math. Comput. Chem.* **67** (2011) 193–212.

[15] H. Hong, B. S. Hernandez, J. Kim, J. K. Kim, Computational translation framework identifies biochemical reaction networks with special topologies and their long-term dynamics, *SIAM J. Appl. Math.* **83** (2023) 1025–1048.

[16] H. F. Fariñas, E. R. Mendoza, A. R. Lao, Chemical reaction network decompositions and realizations of S-systems, *arXiv:2003.01503*.

[17] J. Anderies, S. Carpenter, W. Steffen, J. Rockström, The topology of non-linear global carbon dynamics: from tipping points to planetary boundaries, *Environ. Res. Lett.* **8** (2013) 044–048.

[18] J. B. Rinon, R. Mendoza, V. M. Mendoza, Parameter estimation of an S-system model using hybrid genetic algorithm with the aid of sensitivity analysis, in: P. Fernandez, H. Adorna, A. Sioson (Eds.), *Proceedings of the 19th Philippine Computing Science Congress, National University, City of Manila*, Comput. Soc. Philippines, 2019, pp. 94–102.

[19] M. Feinberg, Chemical reaction network structure and the stability of complex isothermal reactors I: The deficiency zero and deficiency one theorems, *Chem. Eng. Sci.* **42** (1987) 2229–2268.

[20] M. Feinberg, *Foundations of Chemical Reaction Network Theory*, Springer, Cham, 2019.

[21] M. Feinberg, *Lecture on Chemical Reaction Networks*, Univ. of Wisconsin, 1979.

[22] M. D. Johnston, D. Siegel, G. Szederkenyi, A linear programming approach to weak reversibility and linear conjugacy of chemical reaction networks, *J. Math. Chem.* **50** (2011) 274–288.

[23] N. Fortun, A. Lao, L. Razon, E. Mendoza, A deficiency-one algorithm for a power law kinetics with reactant determined interactions, *J. Math. Chem.* **56** (2018) 2929–2962.

[24] N. Fortun, E. Mendoza, Absolute concentration robustness in power law kinetic systems, *MATCH Commun. Math. Comput. Chem.* **85** (2020) 669–691.

[25] N. Fortun, E. Mendoza, Comparative analysis of carbon cycle models via kinetic representations. *J. Math. Chem.* **61** (2023) 896–932.

[26] N. Fortun, E. Mendoza, L. Razon, A. Lao, A deficiency zero theorem for a class of power-law kinetic systems with non-reactant-determined interactions, *MATCH Commun. Math. Comput. Chem.* **81** (2019) 621–638.

[27] O. R. Antonio Stoichiometric matrix, in: W. Dubitzky, O. Wolkenhauer, K. H. Cho, H. Yokota (Eds.), *Encyclopedia of Systems Biology*, Springer, New York, 2013, pp 2014–2014.

[28] R. Curto, A. Sorribas, M. Cascante, Comparative characterization of the fermentation pathway of Saccharomyces cerevisiae using biochemical systems theory and metabolic control analysis: Model definition and nomenclature, *Math. Biosci.* **130** (1995) 25–50.

[29] R. E. Ferner, J. K. Aronson, Cato Guldberg and Peter Waage, the history of the law of mass action, and its relevance to clinical pharmacology, *Br. J. Clin. Pharmacol.* **81** (2016) 52–55.

[30] S. Müller, G. Regensburger, Generalized mass action systems and positive solutions of polynomial equations with real and symbolic exponents, in: V. P. Gerdt, W. Koepf, W. M. Seiler, E. H. Vorozhtsov (Eds.), *Computer Algebra in Scientific Computing*, Springer, Cham, 2014, pp. 302–323.

[31] S. Müller, G. Regensburger, Generalized mass action systems: complex balancing equilibria and sign vectors of the stoichiometric and kinetic order subspaces, *SIAM J. Appl. Math.* **72** (2012) 1926–1947.

[32] V. Heck, J. Donges, W. Hucht, Collateral transgression of planetary boundaries due to climate engineering by terrestrial carbon dioxide removal. *Earth Sys. Dyn.* **7** (2016) 783–796.

A Notation and acronym

We list some of the symbols and acronyms used in the paper.

Table 3. List of Symbols

$\mathscr{S}$	set of species
m	cardinality of species
$\mathscr{C}$	set of complexes
n	cardinality of complexes
$\mathscr{R}$	set of reactions
r	cardinality of reactions
$\mathscr{L}$	set of linkage classes
l	number of linkage classes
sl	number of strong linkage classes
n_r	number of reactant complexes
t	number of terminal strong linkage classes
δ	Deficiency
$\hat{\delta}$	Kinetic reactant deficiency
S	Stoichiometric subspace
$\mathscr{N}$	Reaction network
N	Stoichiometric matrix
K	Kinetics of a CRN
I_K	Interaction map
A_K	Laplacian map
I_a	Incidence matrix
F	Kinetic order matrix
Y	Matrix of complexes
$\hat{T}$	augmented T-matrix
T_κ	poly T-matrix
$\tilde{S}_R$	Kinetic reactant flux subspace
Q	Positive kinetic reactant flux class
$E_+(\mathscr{N}, K)$	Set of positive equilibria
$Z_+(\mathscr{N}, K)$	Set of complex balanced equilibria
$\mathbb{R}$	Set of real numbers

Table 4. Abbreviations

ACR	Absolute concentration robustness
BCR	Balanced concentration robustness
BST	Biochemical systems theory
ILC	Independent linkage classes
MAK	Mass action kinetics
GMA	Generalized mass action
ODE	Ordinary differential equations
PYK	Poly-PL kinetics
TIK	$\hat{T}$-rank maximal kinetics
PL	Power law
RDK	Reactant-determined kinetics
NDK	Non-reactant-determined kinetics
CRN	Chemical reaction network
CKS	Chemical kinetic system
SFRF	Species formation rate functions
ZDD	Zero deficiency decomposition
SF-pair/SF-type	Shinar-Feinberg pair/Shinar-Feinberg type