GRAPH-BASED, DYNAMICS-PRESERVING REDUCTIONS OF (BIO)CHEMICAL SYSTEMS

TALMON SOARES
Bachelor of Science, University of California, Irvine, 2018

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

BIOCHEMISTRY

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

GRAPH-BASED, DYNAMICS-PRESERVING REDUCTIONS OF (BIO)CHEMICAL SYSTEMS

TALMON SOARES

Date of Defence: April 18, 2023

Dr. Marc R. Roussel Thesis Supervisor	Professor	Ph.D.
Dr. Shahadat Hossain Thesis Examination Committee Member	Professor	Ph.D.
Dr. Athan Zovoilis Thesis Examination Committee Member	Associate Professor	Ph.D.
Gheorghe Craciun External Examiner University of Wisconsin-Madison	Professor	Ph.D.
Michael Gerken Chair, Thesis Examination Committee	Professor	Ph.D.

Dedication

To my father, for always telling me that lack of knowledge and experience is not an excuse for not being able to accomplish something. And to my wife, without whom I would probably have never visited Canada.

Abstract

Complex dynamical systems often contain many unknown parameters and variables that may or may not serve as contributors to interesting behaviors a system may exhibit. For chemical and biochemical systems, which are typically quite large, these include (but are not limited to) behaviors stemming from bifurcations such as oscillations, patterns and multistationarity (multiple steady states). Due to the size and complexity of these systems, a dynamics-preserving reduction scheme that is able to isolate the necessary contributors to these systems to not only reduce their complexity, but to also reduce the level of uncertainty a system may have—such as unknown parameters and variables—is desired. The purpose of this thesis is to develop alternative reduction methods for (bio)chemical systems that are modeled by mass-action kinetics that are unlike other common techniques that exploit timescales in stiff models or that are optimization-based. I instead look to a graph-based approach by representing the model as a bipartite graph and investigating its subnetworks known as fragments, which correspond directly to terms in the characteristic polynomial. In this representation, I preserve key elements of these bipartite graphs—*critical fragments*—in order to maintain necessary conditions for behaviors such as positive-feedback oscillations and multistationarity. These results are then applied to an existing model for the transcriptional control of Hmp, an NO detoxifying enzyme, by the iron-sulfur protein FNR that displays bistability. The initial model consists of 15 mass-action reactions and 11 species, and the reduced model ends up with 10 reactions and 7 species.

Contribution of Authors

Chapter 3 is from a draft manuscript written by Marc R. Roussel and myself.

MR Roussel: Conceptualization; Validation; Resources; Writing — Review & Editing; Supervision; Funding acquisition. **T Soares:** Formal analysis; Investigation; Writing — Original Draft; Visualization.

Acknowledgments

First and foremost, I would like to thank Marc R. Roussel, my thesis supervisor, for granting me the opportunity to enroll in an M.Sc. program here in Lethbridge, Alberta. Being from the United States with an undergraduate mathematics degree made enrolling into a graduate biochemistry program in Canada seem like a pipe dream. However, thanks to your professionalism, guidance, and sincerity, that dream came to fruition and quickly became reality. I could not have asked for a better, more caring supervisor. Thank you for motivating me throughout these two years (as well as incorporating some Canadian history and culture here and there)!

I would also like to thank my undergraduate Math Biology professor, Dr. German Enciso, as well as the TA for the course, Daniel Bergman, at the University of California, Irvine for introducing me to chemical kinetics. In addition, I want to thank my undergraduate PDEs professor, Dr. Natalia Komarova, also at the University of California, Irvine for broadening my knowledge of differential equations. Without all of you, I would not have found an application of mathematics that I truly enjoy. Thank you to all for being my academic references for this program.

Last but not least, I want to thank my thesis committee members and the Roussel lab group. Thank you for the kind words of encouragement and all the thought-provoking conversations. Most certainly, thank you for a great time and a fun learning experience.

Contents

List of Figures

Chapter 1

Introduction

Chemical and biochemical systems are (often large) dynamical systems that are capable of exhibiting complex behaviors such as oscillations, patterns and multistationarity (multiple steady states) [1]. These interesting behaviors not only motivate analysis with mathematical modeling, but are also critical to many biological processes: oscillations play an important role in early embryonic development [2]; Turing patterns have often been the basis for explanations in developmental patterning [3, 4]; and multistationarity can account for epigenetic differences [5]. While dynamically distinct, all of these behaviors arise from bifurcations. Bifurcations occur when the stability of a steady state changes. In a closed chemical system, the steady state is tantamount to the equilibrium point, which is always stable [6]. Open systems, however, do not have this limitation and can have multiple steady states that are either stable or unstable [7]. In particular, these behaviors stem from instabilities and thus cannot be exhibited by closed systems near their equilibrium point, but can be exhibited near the steady state(s) of an open system.

One method of analyzing these behaviors is by transforming the set of elementary reactions that constitute the system into a system of ordinary differential equations (ODEs) through the use of the law of mass action. In this manner, we can analyze the system algebraically through several methods, but I will primarily focus on linear stability analysis. With linear stability analysis, we analyze an expression known as the characteristic polynomial that is derived from the Jacobian matrix of the ODEs evaluated at the steady state. The coefficients that appear in this polynomial depend on parameters such as rate constants

and species concentrations. However, the computational work required to execute the algorithms that determine these coefficients can grow very rapidly due to the combinatorial nature of the calculations [8].

To circumvent the computational costs of algebraic methods, graphical methods have been developed that transform the problem of analyzing matrices and their sub-matrices into one that analyzes the properties of graphs and their subgraphs [9–15]. These methods determine the necessary components (species and/or reactions) for a particular behavior to be exhibited by the system. This enables the potential removal of species and/or reactions by deleting species and/or reactions that do not play critical roles in the dynamics of the system. This idea will be revisited later in this thesis where I will develop *conservative* model reduction rules based on a bipartite graph representation of a (bio)chemical reaction mechanism originally introduced for stability analysis by Ivanova [10, 14]. By conservative reductions, I mean a set of graph operations that retain key elements of the structure of the characteristic polynomial so that the potential for bifurcations of the original model is retained. This thesis builds on Ivanova's graphical method to develop dynamics-preserving, graph-based reductions that can be applied to general mass-action models that isolate and preserve the key contributors to dynamics.

In order to prepare the reader, Ch. 2 begins with some of the basic ideas associated with nonlinear dynamics. The majority the chapter focuses on bipartite graphs and their structure, and how fragments (subnetworks of bipartite graphs) correspond to terms in the characteristic polynomial. I end this chapter by defining several terms to be used in my analysis and reduction of mechanisms via their bipartite graphs. In Ch. 3, I present the main ideas of this thesis that serve as the basis for model reduction, and both develop and prove the theorems that enable these reductions. The chapter ends with the application of these reductions to a model for the control of the synthesis of the NO-detoxifying enzyme Hmp. Lastly, Ch. 4 concludes this thesis with a discussion of the disadvantages, advantages and possible extensions of these graph-based reductions.

Chapter 2

Background

2.1 Nonlinear Dynamics

2.1.1 Law of Mass Action, Steady States & Linearization

Chemical kinetics is the study of how fast chemical reactions proceed from reactants to products and is illustrative of the connection between chemistry and mathematics [16]. In this section, we will focus on (bio)chemical systems modeled by mass-action kinetics and recall methods from [14, 17]. A (bio)chemical mechanism with m chemical substances A_k, p chemical reactions and positive rate constants k_i can be written as

$$\sum_{k=1}^{m} \alpha_{ik} \mathrm{A}_k \xrightarrow{k_i} \sum_{k=1}^{m} \beta_{ik} \mathrm{A}_k, \quad i = 1, ..., p. \tag{2.1}$$

The constants α_{ik} and β_{ik} are non-negative integers called stoichiometric coefficients. Let u_k represent the concentration of A_k and $\mathbf{u}$ be the vector of concentrations. Then the Law of Mass Action gives the rate of the ith reaction as

$$w_i(\mathbf{u}) = k_i u_1^{\alpha_{i1}} \cdots u_m^{\alpha_{im}}, \quad i = 1, ..., p. \tag{2.2}$$

Then, by using Eqs. (2.1) and (2.2), we have the following ODE system that models the concentrations of each species over time:

$$\frac{du_k}{dt} = \sum_{i=1}^{p} (\beta_{ik} - \alpha_{ik}) w_i(\mathbf{u}) = f_k(\mathbf{u}), \quad k = 1, ..., m. \tag{2.3}$$

If $\mathbf{u}$ is a steady state, it satisfies $f_k(\mathbf{u}) = 0$, $\forall k \in [1, m]$. Often, (bio)chemical systems are nonlinear, and one way of analyzing the stability of steady states is to use linear stability analysis. We use a first-order multivariate Taylor expansion to linearize the system around its steady state. This expansion generates the Jacobian, which is given by

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial u_1} & \cdots & \frac{\partial f_m}{\partial u_m} \end{bmatrix}. \tag{2.4}$$

J is always square with m rows and m columns. We then evaluate J at the steady state.

2.1.2 Eigenvalues and Eigenvectors

The linearized, homogenous first-order ODE system in matrix form is given by

$$\mathbf{u}' = J\mathbf{u}, \tag{2.5}$$

where $\mathbf{u}'$ is the time derivative of $\mathbf{u}$, and will have general solutions of different forms, depending on the rank of J. For example, if the rank of J is full (m), the system's general solution will be

$$\mathbf{u} = \sum_{i=1}^{m} c_i e^{\lambda_i t} \mathbf{v}_i, \tag{2.6}$$

where λ_i is the eigenvalue associated with the eigenvector $\mathbf{v}_i$, and c_i is an arbitrary constant. If J is rank-deficient, the general solution would contain some orders of t with the terms associated with the repeated eigenvalues. Regardless of the form, its eigenvalues will always appear in the solution's exponent terms. A λ_i with a positive real part would cause its corresponding term in the solution to diverge to infinity for large t, whereas a λ_i with a negative real part would cause its corresponding term to converge to zero. Hence, if the eigenvalues of the Jacobian evaluated at the steady state contain at least one eigenvalue whose real part is positive, then the steady state is unstable as it would cause the solution

to diverge to infinity in the linearized form. In the original model, this translates to the trajectories diverging from the vicinity of the steady state. Thus, the stability of a steady state is determined by its eigenvalues.[1] So how do we find these eigenvalues?

2.1.3 The Characteristic Polynomial

Eigenvalues can be determined by finding the zeroes of the characteristic polynomial of the Jacobian evaluated at a steady state. The characteristic equation arising from the stability analysis of a steady state of an ODE is $\det(J - \lambda I) = 0$, where I is the $m \times m$ identity matrix, or

$$p(\lambda) = \lambda^{m-n}(\lambda^n + a_1\lambda^{n-1} + \ldots + a_b\lambda^{n-b} + \ldots + a_n) = 0, \tag{2.7}$$

where m is the size of J given by the number of species, n is the number of *independent* species, and $m-n$ is the number of conservation relations of the system [19]. The rank of J is equal to the number of independent species in the (bio)chemical system. It is important to write the polynomial in this manner as the necessary conditions for complex behaviors (discussed in a later section) relate to the number of independent species in a system and not to the total number of species.

2.2 Methodology

2.2.1 The Characteristic Polynomial from Algebraic Methods

To set the basis for graphical methods, we first look at the algebraic methods to both relate our graphical methods and motivate the need for a simpler scheme. In this section, we recall coefficient formulas for the characteristic polynomial adapted from [8]. The coefficients of Eq. (2.7) can be found by taking the trace and several combinations of principal

[1] It is important to note that eigenvalues with a zero real part would require a nonlinear theory known as center manifolds and will not be discussed here as our focus is on linear stability analysis [18].

minors of J and are given by

$$\begin{aligned} a_1 &= (-1)^1 \cdot \sum_{i=1}^{m} J_{ii} = (-1) \cdot \mathrm{Tr}(J), \\ a_b &= (-1)^b \cdot \sum_{i=1}^{l} M_i, \quad 2 \leq b \in \mathbb{N} \leq r. \end{aligned} \tag{2.8}$$

a_1 is straightforward as it is the negative of the trace of J. a_b is given by the sum of all $b \times b$ principal minors, M_i. A principal minor is the determinant of a square submatrix formed by the intersections of rows and columns of the same indices. l refers to the number of combinations of b columns/rows chosen from the m possibilities. To illustrate how to find a_b, consider a 3×3 matrix with rank equal to 2. The following is its characteristic polynomial:

$$p(\lambda) = \lambda(\lambda^2 + a_1\lambda^1 + a_2\lambda^0), \tag{2.9}$$

where

$$\begin{aligned} a_1 &= -\mathrm{Tr}(J), \\ a_2 &= (-1)^2 \cdot (M_{1,2} + M_{1,3} + M_{2,3}). \end{aligned} \tag{2.10}$$

$M_{1,2}$, for example, is read as the determinant of the 2×2 matrix formed by the intersection of rows 1 and 2 and columns 1 and 2. Here, a_2 means b is 2, so a_2 requires all the combinations of 2×2 principal minors of the 3×3 matrix J, giving the three minors in Eq. (2.10). There are three minors to compute since the number of combinations of $b = 2$ in $m = 3$ is $\binom{3}{2} = 3$. Due to the coefficients being combinatorial in nature, the different minors that need to be computed in larger systems can grow very rapidly.[2] Thus, we want to employ a method that can avoid the combinatorial blow-up we may have from algebraic methods.

2.2.2 Graphical Methods - The Bipartite Graph

A bipartite graph consists of two non-intersecting sets of vertices representing the species and reactions of the system and directed edges starting from one type of vertex and ending

[2] For example, coefficients a_4 and a_5 of a system with 9 independent species would contain 126 different minors of orders 4 and 5 to compute since $\binom{9}{4} = \binom{9}{5} = 126$.

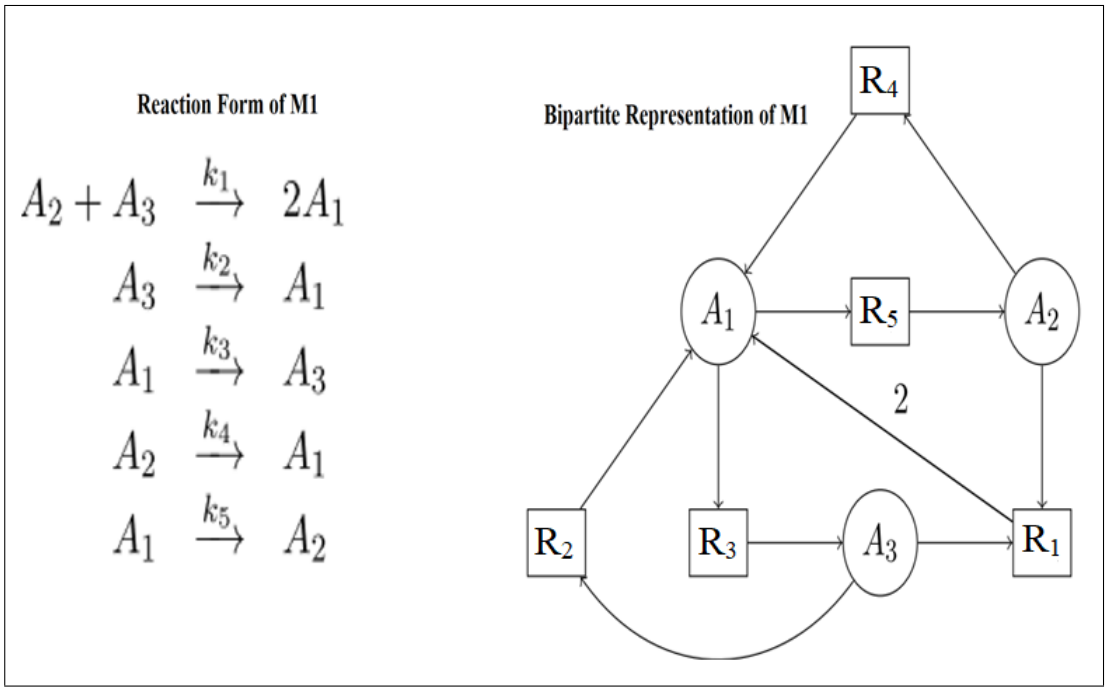

Figure 2.1: This mechanism contains 3 species and 5 reactions and has one mass-conservation relation, and thus 2 independent species. Adapted from [19].

at the other type of vertex [19]. We will use [A_1...A_m] to represent the species vertices and [R_1...R_p] for the reaction vertices. It is important to note that in our labeling of the bipartite graph, R_j will also represent the reaction rate of its corresponding reaction. These connected vertices form elements known as edges, paths, cycles, fragments and subgraphs of fragments. (See Sect. 2.3.1 for their definitions.) Fragments, simply put, are subnetworks of bipartite graphs and serve as graphical representations of terms found in the characteristic polynomial. Figure 2.1 depicts a chemical mechanism and its bipartite graph from [19] named M1 for convenience. In this graph, circles represent species vertices and squares represent reaction vertices. M1 contains three species with only two independent species, making the rank of its stoichiometric matrix two. Figures 2.2 and 2.3 show a few elements of M1 to illustrate what edges, paths and cycles may look like in a bipartite graph.

With a few examples of how to read bipartite graphs and their elements illustrated in Figs. 2.1–2.3, we can discuss their most important components: fragments. Fragments are important because they correspond to terms in the characteristic polynomial. Let $\{A_z\}$ and

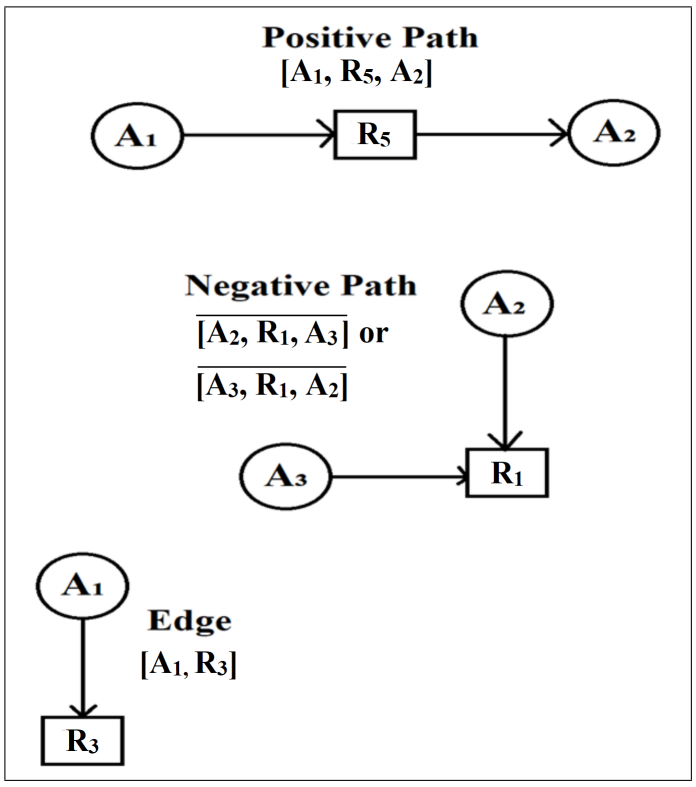

Figure 2.2: Examples of the different types of paths and an edge, along with their notation, that can be found in M1.

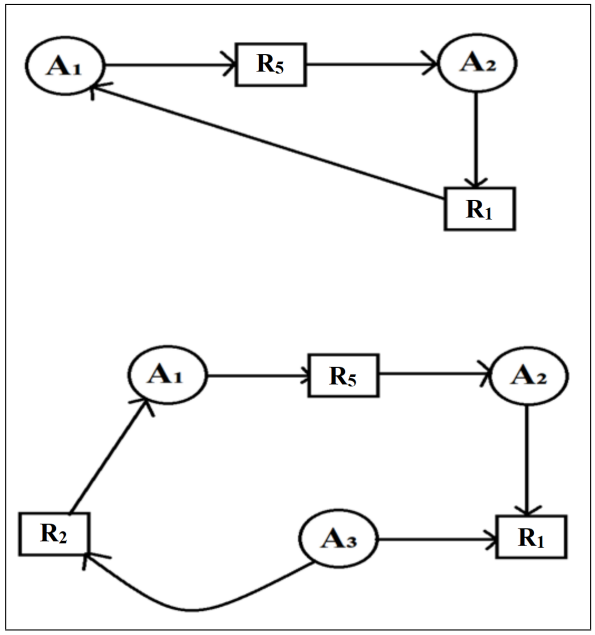

Figure 2.3: Examples of positive (top) and negative (bottom) cycles in M1.

$\{R_h\}$ be sets of i different species and j different reaction rates, respectively, in a fragment. Then, this fragment corresponds to a term in the characteristic polynomial of the form

$$(\prod_{z=1}^{i}\prod_{h=1}^{j}\frac{R_h}{A_z})\lambda^{n-i} \tag{2.11}$$

where n is the number of independent species in the system and i is the order of the fragment, which is given by its number of species. In short, a fragment corresponds to a term in the coefficient of the λ^{n-i} term in the characteristic polynomial, consisting of the product of the fragment's j reaction rates over the product of its i species concentrations. From [14], each fragment S_k and its corresponding term also have a coefficient, K_{S_k}, given by

$$K_{S_k} = \sum_{g \in S_k} K_g, \tag{2.12}$$

where

$$K_g = (-1)^{t_g} \prod_{[\text{edges} \in g]} (\alpha_{jk})^2 \prod_{[\text{cycles} \in g]} K_C, \tag{2.13a}$$

$$K_C = \prod_{[\text{negative paths} \in C]} (-\alpha_{jk}\alpha_{ji}) \prod_{[\text{positive paths} \in C]} \alpha_{jk}\beta_{ji}. \tag{2.13b}$$

In these equations, g labels subgraphs, C labels cycles, t_g is the number of cycles in a subgraph, and α_{jk} and β_{ji} are the same stoichiometric coefficients from Eq. (2.1). The edges considered in a fragment are reactant to reaction edges. Paths are reactant-reaction-species sequences with their connecting edges, with positive paths encoding reactant-product relationships, and negative paths encoding co-reactant relationships. A cycle is composed of paths. Eq. (2.12) gives a fragment's coefficient as the sum of each its subgraphs' coefficients, K_g, which are given in turn by Eqs. (2.13).

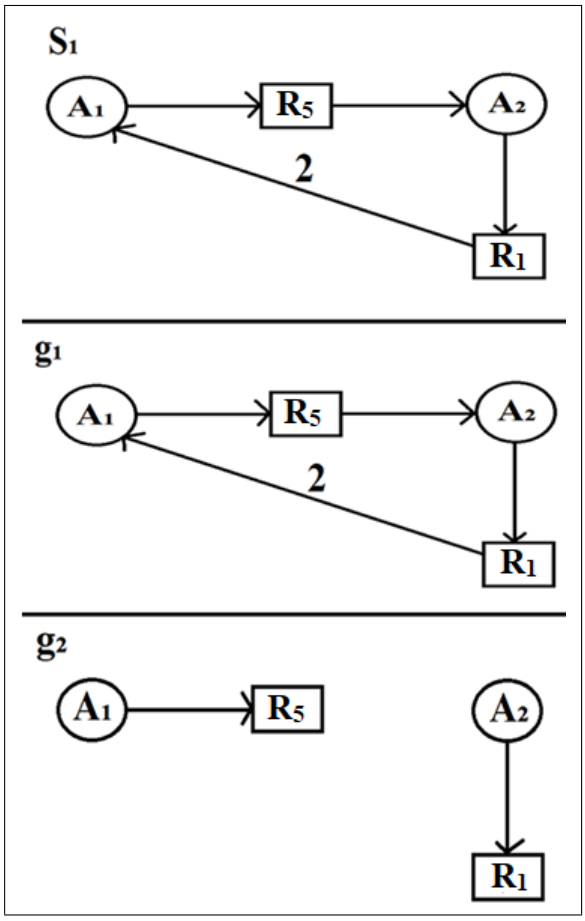

Figure 2.4: A fragment of M1, denoted by S_1, and its two subgraphs g_1 and g_2.

2.2.3 Illustrating the Connection

Fragment S_1 in Fig. 2.4 corresponds to the term in the characteristic polynomial of the form $\frac{R_1R_5}{A_1A_2}\lambda^0$ since the fragment's order is 2 and M1 contains two independent species. To calculate its numerical coefficient, we use Eqs. (2.12) and (2.13) to get a value of -1.[3] This means that the polynomial will have a $-\frac{R_1R_5}{A_1A_2}$ term in its constant (λ^0) expression. From [19], M1 has the following ODE system:

$$\begin{aligned}
\dot{A_1} &= 2k_1A_2A_3 + k_2A_3 - k_3A_1 + k_4A_2 - k_5A_1, \\
\dot{A_2} &= -k_1A_2A_3 - k_4A_2 + k_5A_1, \\
\dot{A_3} &= -k_1A_2A_3 - k_2A_3 + k_3A_1.
\end{aligned} \tag{2.14}$$

[3]Since g_1 contains one cycle and only positive paths, $K_{g_1} = (-1)^1 \cdot 1 \cdot 2 = -2$. The second subgraph, g_2, only consists of edges, so $K_{g_2} = (-1)^0 \cdot 1^2 \cdot 1^2 = 1$. Thus, $K_{S_1} = K_{g_1} + K_{g_2} = -2 + 1 = -1$.

Note the conservation relation, $A_1 + A_2 + A_3 = A_{\text{tot}}$. The reaction rates are given by:

$$\begin{aligned} R_1 &= k_1 A_2 A_3, \\ R_2 &= k_2 A_3, \\ R_3 &= k_3 A_1, \\ R_4 &= k_4 A_2, \\ R_5 &= k_5 A_1. \end{aligned} \tag{2.15}$$

Using Eqs. (2.4), (2.14) and (2.15), we get the following Jacobian [19]:

$$J = \begin{bmatrix} -\frac{R_3}{A_1} - \frac{R_5}{A_1} & \frac{2R_1}{A_2} + \frac{R_4}{A_2} & 2\frac{R_1}{A_3} + \frac{R_2}{A_3} \\ \frac{R_5}{A_1} & -\frac{R_1}{A_2} - \frac{R_4}{A_2} & -\frac{R_1}{A_3} \\ \frac{R_3}{A_1} & -\frac{R_1}{A_2} & -\frac{R_1}{A_3} - \frac{R_2}{A_3} \end{bmatrix} \tag{2.16}$$

Since the characteristic polynomial of M1 is of the forms of Eqs. (2.9) and (2.10), we have from [19] that

$$a_2 = \frac{R_1R_3 - R_1R_5 + R_3R_4}{A_1A_2} + \frac{-R_1R_3 + R_1R_5 + R_2R_5}{A_1A_3} + \frac{R_1R_2 + R_1R_4 + R_2R_4}{A_2A_3}. \tag{2.17}$$

The term in red is *exactly* the term that fragment S_1 corresponds to, with its associated coefficient of -1. This method is quite useful as it allows us to determine the coefficients of the characteristic polynomial without having to take any determinants. But, why are these coefficients important? The answer stems from the sign of the coefficients of these fragments, as we will see next.

2.2.4 Critical Fragment Theory

When a (bio)chemical system is represented as a bipartite graph, we can apply what is known as critical fragment theory. When the coefficient of a fragment is negative, we call it a **critical fragment** [14, 17, 19, 20]. The appearance of a critical fragment in the bipartite

graph of a chemical mechanism is *necessary* for certain instabilities to appear. In particular, we have the following two results [14, 19]:

- A critical fragment of order equal to the number of independent species, n, is necessary for the possibility of a saddle-node bifurcation, which would allow for multistationarity.[4]
- A critical fragment or order less than n is necessary for oscillations due to positive feedback.[5]

These are powerful conditions as they enable a determination of the potential for multistationarity or positive-feedback oscillations based on the existence (or lack thereof) of critical fragments [20]. It is important to note that these are *necessary* and not sufficient conditions. While unfortunately not enough to guarantee certain behaviors in these systems, these do serve as a basis for the reduction methods in this thesis. The main idea here is that we must preserve these critical fragments through any reductions applied. Ideally, this means that preserving critical fragments also preserves dynamics. Before we delve further into reductions via this graphical approach, key terms and their definitions (as well as notations) are outlined in the next section.

2.3 Notation and Definitions

2.3.1 Canonical Terms

This section outlines the terminology established by Mincheva and Roussel [14] based on the English translation of Ivanova's original Russian terminology [10].

(Reactant) Edge — A directed arrow from a reactant species vertex A_k to a reaction vertex R_j, denoted $[\mathrm{A}_k, \mathrm{R}_j]$. These edges are the only kind specifically named in the classic version of critical fragment theory.

[4] Note that a critical fragment of this order would correspond to the constant term in the characteristic polynomial. This means that the constant term *must* contain negative terms!

[5] This would correspond to a term in a coefficient a_k of the characteristic polynomial with $k < n$. In other words, we require at least one term in some a_k, $k < n$, to be negative.

Product Edge — A directed arrow from a reaction vertex R_k to a product species vertex A_i, denoted $[\mathrm{R}_k, \mathrm{A}_i]$.

Edge weight — The weight of an edge (reactant or product) is the corresponding stoichiometric coefficient.

Positive path — A sequence of directed arrows from a reactant species vertex to a reaction vertex and then from that same reaction vertex to a product species vertex, denoted $[\mathrm{A}_k, \mathrm{R}_j, \mathrm{A}_i]$.

Negative path — A sequence of directed arrows from a reactant species vertex to a reaction vertex, followed by an edge traversed against the direction of its arrow to another reactant species vertex, denoted $\overline{[\mathrm{A}_k, \mathrm{R}_j, \mathrm{A}_i]}$.

Cycle — A sequence of paths that start and end at the same species vertex. A cycle has a positive parity if it contains an even number of negative paths, and a negative parity if it contains an odd number of negative paths.

Subgraph — A union of edges and cycles in which each species is at the origin of only one edge or path. Each subgraph has a coefficient, K_g, computed by Eq. (2.13). A subgraph's coefficient is thus determined by the stoichiometric coefficients appearing in its edges and cycles, and the number and parities of cycles.

Fragment — A union of subgraphs that share the same vertices. Each fragment corresponds to a term in the characteristic polynomial of the form (2.11).

Coefficient of a fragment — The coefficient of a fragment, K_{S_k}, is given by the sum of all of its subgraphs' coefficients [Eq. (2.12)].

Critical fragment — A fragment with a negative coefficient.

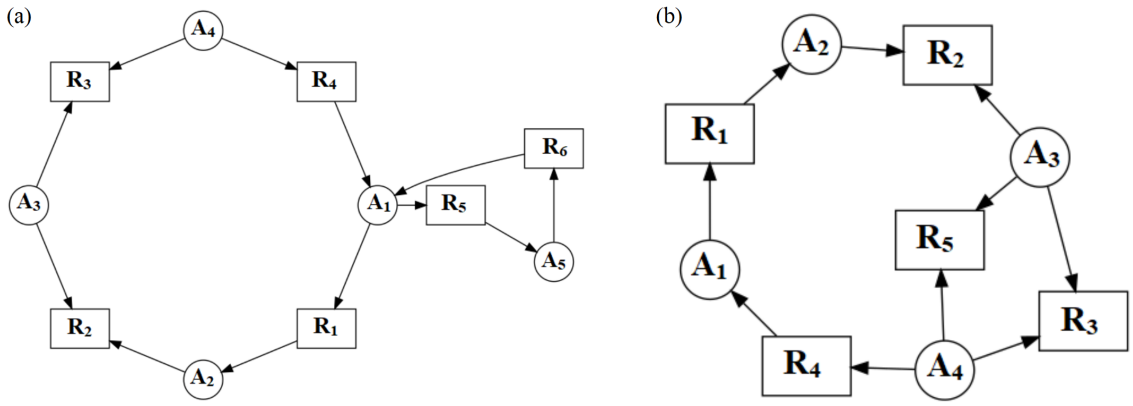

Figure 2.5: Examples of linked cycles. (a) A basic form of a linked cycle. (b) A slightly more complicated linked cycle, where more than one vertex is involved in another cycle.

2.3.2 (Non-Canonical) Reduction Terms

This section outlines new terminology for bipartite graphs developed for the purpose of model reduction via the graphical approach.

Isolated cycle — A cycle C_i whose vertices are not involved in any other cycle. This means that if the vertices of C_i and all of their associated edges form a fragment, then no cycle other than C_i appears in a subgraph.

Linked cycle — A cycle C_L where at least one vertex in C_L is part of another cycle that is not C_L. Fig. 2.5 shows examples of linked cycles.

Unidirectional cycle — A cycle that contains at least one positive path. As a consequence of the positive path, unidirectional cycles can only be traversed in one direction.

Entry vertex — A vertex within a cycle C that contains at least one arrow connected to it that is not part of C.

Isolated edge — An edge whose vertices are not entry vertices.

Integrated cycle — A cycle with one or more entry vertices. See Fig. 2.6 for examples of isolated integrated cycles. Note that linked cycles are necessarily integrated.

Unitary edge — An edge (reactant or product) whose weight is one.

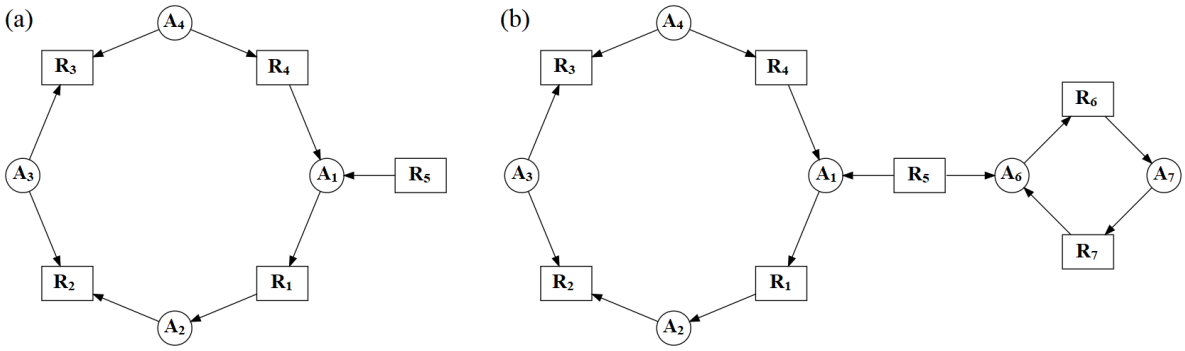

Figure 2.6: Examples of integrated isolated cycles. (a) An integrated isolated cycle with a single entry vertex A_1. (b) Two separate integrated isolated cycles. The cycle on the left has the species vertex A_1 as its entry vertex where the cycle on the right has the species vertex A_6 as its entry vertex. Both of these cycles are considered isolated cycles since none of their vertices are part of other cycles.

Unitary (positive/negative) path — A path (positive or negative) consisting of unitary edges.

Unitary cycle — A cycle consisting entirely of unitary paths.

2.3.3 Notation

- $[A_k, R_j, A_i]_{a_h, a_{h+1}}$ represents the positive path $[A_k, R_j, A_i]$ with arrows of weight a_h leaving A_k and weight a_{h+1} leaving R_j.

- $\overline{[A_k, R_j, A_i]}_{a_h, a_{h+1}}$ represents the negative path $\overline{[A_k, R_j, A_i]}$ with arrows of weight a_h leaving A_k and weight a_{h+1} leaving A_i.

Chapter 3

Reducing (Bio)chemical Systems

Chapter 3 is from a draft manuscript written by Marc R. Roussel and myself.

CRediT authorship contribution statement

MR Roussel: Conceptualization; Validation; Resources; Writing — Review & Editing; Supervision; Funding acquisition. **T Soares:** Formal analysis; Investigation; Writing — Original Draft; Visualization.

3.1 Current Methods

Model reduction methods for (bio)chemical systems have recently been reviewed by Snowden and coworkers [21] and by Gorban [22]. Timescale exploitation methods are the most commonly applied approaches for reducing models of (bio)chemical systems [21,23]. This means that the reactions in these systems often occur across a wide range of timescales, where "fast" variables can be disregarded when analyzing the "slow" variables. Other methods that do not involve timescale separation such as some optimization-based approaches [24], techniques using sensitivity analysis to eliminate parameters and potentially chemical species or reactions [25] and lumping [26] exist but also carry with them their own limitations. For example, most of these methods of model reduction require parameter values to be known and have the potential to be computationally intensive. We seek to reduce large models from a different, graphical approach to not only aid in alleviating some of the computation, but also to add another tool to the tool box of reduction techniques. One great advantage graph-theoretical methods have over more conventional reduction methods

is that the necessary conditions for bifurcations can be decided in the absence of parameter values. Our goals with graph-theoretical methods are to identify system dynamics and to preserve these dynamics through graph-based reductions.

3.2 A Graphical Approach — Analyzing Critical Fragments

As mentioned previously, we will use critical fragment theory as a basis for model reduction. The basic concept is simple: since critical fragments are necessary for behaviors such as oscillations due to positive feedback and multistationarity, then preservation of these fragments is imperative. Let us examine what is required for a fragment to be critical.

The coefficient of a fragment, given by Eq. (2.12), is determined as the sum of its subgraphs' coefficients, where a subgraph's coefficient is given by Eq. (2.13); hence, we must examine the structure of a fragment's subgraphs. From Eq. (2.13), we see that a cycle is *necessary* for a negative sign to appear when calculating a fragment's coefficient. Furthermore, positive paths within cycles contribute to positive terms and negative paths contribute to negative terms. Thus, we look to the structure of cycles.

3.2.1 (Unidirectional) Isolated Cycles

Cycles are necessary components of critical fragments, so before any reductions could be applied, we need to know if any changes would affect only the cycle or also other parts of the bipartite graph. This led to the distinction of two different types of cycles: isolated and linked. The main reason for this distinction is the connectivity of the vertices within cycles to other parts of the bipartite graph.

From experimenting with cycles in critical fragments, it became clear that reductions could be applied immediately to cycles that contained vertices that were not part of other cycles. In other words, reductions were more easily implemented in cycles whose vertices were isolated from other cycles—the isolated cycles defined above. Equally significantly, in order to apply conservative reductions, unidirectionality of these cycles must be established

by requiring each isolated cycle to contain at least one positive path. Since a fragment's coefficient depends on the number and types of cycles it contains, reductions involving non-unidirectional cycles are much more tricky since eliminating components of such a cycle will normally change the number of subgraphs in a fragment. In essence, the issue is that non-unidirectional cycles are made of sequences of negative paths, but a negative path and its reverse is a cycle [14]. Pruning out an edge that is part of a negative path thus necessarily reduces the number of cycles that can be made from a larger non-unidirectional cycle. In addition, since cycles will usually be attached to other vertices in the bipartite graph, the vertex (or vertices) that bridges the cycle to other parts of the bipartite graph must be preserved. This type of vertex is called an entry vertex. When a cycle contains an entry vertex, it is called an integrated cycle.

With an isolated cycle, we do not have to worry about the behavior near an entry vertex because changes to these surrounding vertices will not interfere with the vertices outside of this cycle that are attached to any entry vertex. For example, removing a unitary edge along a positive path that does not contain an entry vertex from an isolated cycle would remove a factor of 1 from the coefficients of any subgraphs in which this cycle appears, and thus leave the coefficient of the fragment invariant. This means that if the fragment was (non-)critical, it would remain (non-)critical. Similarly, since a negative path contributes a negative factor, then removing two unitary negative paths not containing an entry vertex from an isolated cycle would remove a product of $-1 \cdot -1 = 1$ from the coefficients of subgraphs in which these paths appear, again leaving the fragment's coefficient invariant.[6] The main idea we pursue below is that any number of unitary edges can be removed, whereas unitary negative paths may be removable in pairs. We will eventually lift the requirement for paths to be unitary, which will require transformations of the original graph beyond the removal of edges or paths.

[6]There are however some important subtleties in the case of removing negative paths to which we will return later.

3.2.2 Linked Cycles

When an entry vertex is involved in another cycle, reductions are unfortunately not as straightforward. Cycles of this type are called linked cycles. Note that, by definition, linked cycles involve at least two cycles. When trying to delete vertices near an entry vertex, creation or destruction of cycles may occur even if the vertices involved were connected by a unitary edge along a positive path.

The coefficient of a fragment is the sum of the coefficients of its subgraphs [Eq. (2.12)]. As a rule, changing the number of subgraphs in a fragment can therefore change the coefficient of the fragment, unless the removals are balanced in the sense that the sum of the coefficients of the removed subgraphs is zero. In general, this is difficult to guarantee if we make a transformation that reduces the number of subgraphs in a fragment. Therefore, to simplify linked cycles, a conservative approach will retain the number of subgraphs that can be drawn as well as their coefficients. The key to accomplishing this lies in the set of unidirectional cycles within the linked cycle. If we can find isolated edges (with vertices that are not entry vertices) in a unidirectional cycle within a linked cycle, then similar reductions to those discussed in Sect. 3.2.1 can be applied if unidirectionality of each cycle in a linked cycle is preserved. In general, an entry vertex cannot be removed and reductions near the entry vertex must be applied carefully, otherwise unidirectionality of cycles may be lost, or the number of subgraphs may change.

Figure 3.1 illustrates a linked cycle which, on the surface, may appear to be isolated. However, due to the entry vertex being a reaction vertex involved in a negative path with the larger cycle and the external vertex, A_5, it becomes a linked cycle. If Fig. 3.1 illustrates a fragment of a bipartite graph, it in fact contains several cycles, each of which appears in at least one subgraph of the fragment:

- the "obvious" cycle $\{[A_1, R_1, A_2], \ldots, [A_4, R_4, A_1]\}$;
- the negative-path cycles $\left\{\overline{[A_2, R_2, A_3]}, \overline{[A_3, R_2, A_2]}\right\}$, $\left\{\overline{[A_3, R_3, A_4]}, \overline{[A_4, R_3, A_3]}\right\}$ and $\left\{\overline{[A_4, R_4, A_5]}, \overline{[A_5, R_4, A_4]}\right\}$;

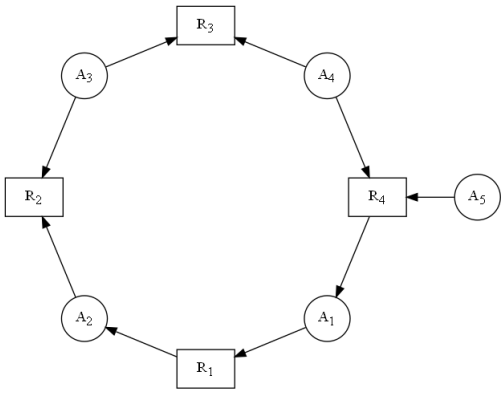

Figure 3.1: A non-obvious linked cycle. On the surface, this cycle may appear to be isolated. However, if these vertices and edges form a fragment, the negative-path cycle $\left\{\overline{[A_4, R_4, A_5]}, \overline{[A_5, R_4, A_4]}\right\}$ (among others) appears in one of the subgraphs, breaking the condition for an isolated cycle.

- the "consecutive negative-path" cycles

$$\left\{\overline{[A_2, R_2, A_3]}, \overline{[A_3, R_3, A_4]}, \overline{[A_4, R_3, A_3]}, \overline{[A_3, R_2, A_2]}\right\},$$
$$\left\{\overline{[A_3, R_3, A_4]}, \overline{[A_4, R_4, A_5]}, \overline{[A_5, R_4, A_4]}, \overline{[A_4, R_3, A_3]}\right\}$$
$$\text{and } \left\{\overline{[A_2, R_2, A_3]}, \overline{[A_3, R_3, A_4]}, \overline{[A_4, R_4, A_5]}, \overline{[A_5, R_4, A_4]}, \overline{[A_4, R_3, A_3]}, \overline{[A_3, R_2, A_2]}\right\};$$

- and a cycle encompassing the entire set of vertices: $\left\{[A_5, R_4, A_1], \ldots, \overline{[A_4, R_4, A_5]}\right\}$.

An attempted simplification through pairwise removal of the negative paths $\overline{[A_2, R_2, A_3]}$ and $\overline{[A_3, R_3, A_4]}$ by deleting vertices A_2, R_2, A_3 and R_3 would drastically reduce the number of subgraphs of the fragment, and might thus alter the subgraph's coefficient. However, the profusion of cycles in subgraphs of this fragment was only possible because of the negative path $\overline{[A_4, R_4, A_5]}$. Note that if A_5 was a product rather than a reactant of R_4, the cycle becomes an isolated cycle and none of these difficulties arise. This example strikes a note of caution with respect to the removal of negative paths when an entry reaction vertex participates in a negative path with vertices that appear in two distinct cycles. Despite the more delicate nature of linked cycles, the same ideas regarding reductions applied to isolated cycles still hold, but we will need to be more careful with the vertex eliminations we allow.

3.3 Graph-Based Reduction Rules

3.3.1 Theorems for Cycles in Bipartite Graphs

In this section, we will present theorems for reduction via the elimination of positive paths, consecutive negative paths, and disjoint negative paths. Figure 3.2 provides a visual aid to the notation used in describing the reductions in the theorems. All of the cycles drawn are unidirectional. They are drawn so that they are traversed clockwise, although this direction is arbitrary. The weight a_h refers to an arbitrarily selected edge in a cycle of order n. By convention, $a_0 \equiv a_n$ and $a_{n+1} \equiv a_1$. Also, the vertices (and their accompanying arrows) are labeled *in the direction of the (unidirectional) cycle*.

Refer to Fig. 3.2(a) for the interpretation of the symbols in Theorem 1.

Theorem 1 (Positive Paths in Unidirectional Isolated Cycles)**.** *Assume we have a bipartite graph that contains an integrated unidirectional isolated cycle C_i of order n, with n sufficiently large.*[7] *Let $a_1, a_2, \ldots, a_n$ be the weights of all n arrows in C_i, with W_o the set containing the weights a_{h-1}, a_{h+1}, and W_e the set containing the weights a_h, a_{h+2}. Let $[A_k, R_j, A_i]_{a_h, a_{h+1}}$ be a positive path in C_i, where neither R_j nor at least one of the species vertices is an entry vertex. Additionally, let R_p and R_s be the reaction vertices immediately preceding and succeeding, respectively, this positive path. We have the following cases:*

1. *A_k is not an entry vertex and at most one element in W_o is non-unitary. Let $\omega_o = \max(W_o)$. If a_h is unitary, then construct the resulting cycle $\overline{C}_i$ by removing species A_k and reaction R_j (and arrows with weights a_{h-1}, a_h, a_{h+1}) and adding a new arrow that connects R_p and A_i with weight ω_o oriented in the same direction as the original arrow with weight a_{h-1}.*

2. *A_i is not an entry vertex and at most one element in W_e is non-unitary. Let $\omega_e = \max(W_e)$. If a_{h+1} is unitary, then construct the resulting cycle $\overline{C}_i$ by removing species*

[7] Exactly how large n needs to be depends on what is to either side of the vertices to be eliminated, notably whether the positive path is adjacent to a negative path.

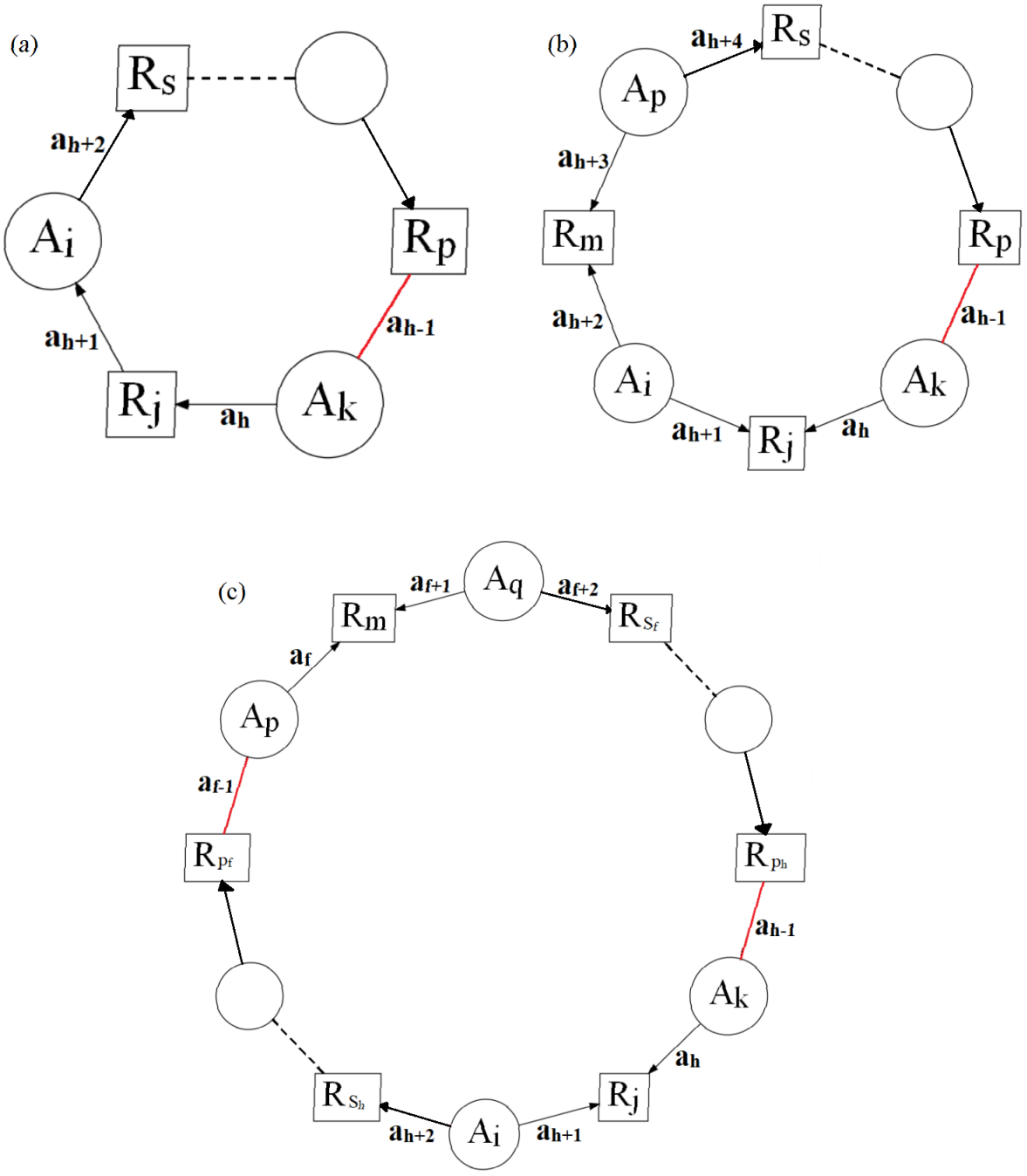

Figure 3.2: Definition of symbols used in describing the reductions of (a) Theorem 1, (b) Theorem 2 and (c) Theorem 3. The red lines represent arrows that can be oriented in either direction. The dotted black lines are there to show there is an arbitrary number of paths in between the vertices. It is *very* important to note that there must be at least one positive path (pre- and post-reduction) oriented in the clockwise direction in each figure in order for reduction to be successfully undertaken.

A_i and reaction R_j (and arrows with weights a_h, a_{h+1}, a_{h+2}) and adding a new arrow that connects A_k and R_s with weight ω_e oriented in the direction of the cycle [clockwise as the cycle is drawn in Fig. 3.2(a)].

If the resulting cycle $\overline{C}_i$ is unidirectional, then any fragment that contained C_i will have the same coefficient coefficient after replacing the latter cycle by $\overline{C}_i$.

Proof. Let C_i be an integrated unidirectional isolated cycle of order n in a bipartite graph. Because of its unidirectionality and isolation, the vertices, paths and edges of C_i can appear in a subgraph in exactly two ways: as the entire cycle, or as a union of reactant edges, in the latter case with all of the edges selected for the subgraph pointing in the direction of the unidirectional cycle. We only allow the removal of non-entry vertices; thus, simplifying the cycle will not affect connectivity to the rest of the bipartite graph. Note in particular that R_j is a non-entry vertex.

In case 1, we have the non-entry vertex A_k with $a_h = 1$. The reactant edge $[\mathrm{A}_k, \mathrm{R}_j]_{a_h}$ contributes a factor of $a_h^2 = 1$ to the coefficient of any subgraph that includes it so removing this edge has no effect on the coefficients of these subgraphs. Furthermore, in any subgraph containing the entire cycle, arrows with weights a_{h-1}, a_h and a_{h+1} contribute a magnitude of $(a_{h-1})(a_h)(a_{h+1}) = \omega_o$ to K_C, and thus to the coefficient of any subgraph that contains C_i. Connecting $\mathrm{R_p}$ and A_i with an arrow of weight ω_o therefore preserves the magnitude of the coefficient of the original cycle for $\overline{C}_i$. Having the new arrow drawn in the same direction as the original arrow between $\mathrm{R_p}$ and A_k preserves the parity of the cycle, and therefore the sign of K_C. Moreover, if unidirectionality was preserved, the cycle resulting from simplification can still appear in a subgraph in only two ways, as a union of reactant edges or as an entire cycle. It follows that the removal of the reactant edge $[\mathrm{A}_k, \mathrm{R}_j]_{a_h}$ with an appropriate arrow drawn between $\mathrm{R_p}$ and A_i preserves the number of subgraphs of any fragments in which the cycle appears as well as the coefficients of the subgraphs, and thus the coefficient of the fragment. Case 2 is proved similarly. □

Figure 3.3 illustrates the use of Theorem 1.

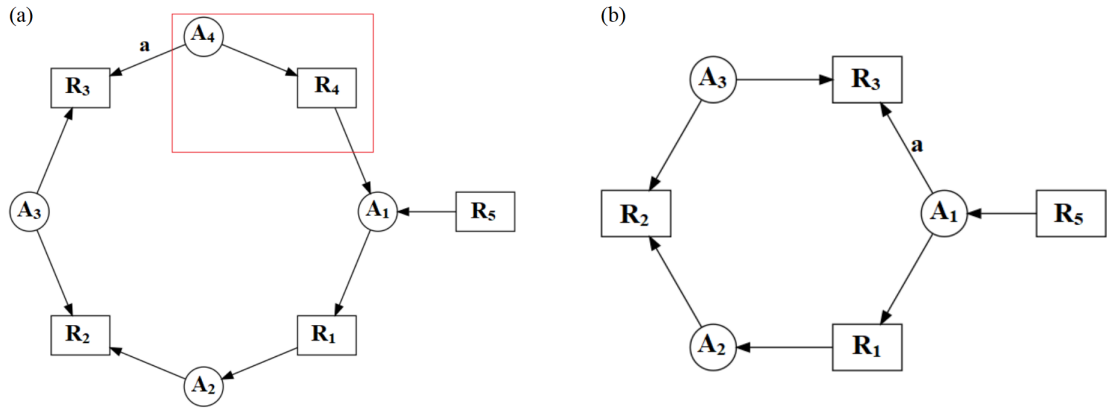

Figure 3.3: (a) An integrated unidirectional isolated cycle that meets the criteria of Theorem 1. We assume that the A_4 to R_3 edge has weight $a_{h-1} = a$. The vertices in the red box are the vertices that will be removed from the positive path [A_4, R_4, A_1]. (b) The result of applying Theorem 1. Note that no further edges could be removed using Theorem 1 because doing so would result in a cycle that was no longer be unidirectional, violating one of the conditions of the theorem.

Remark 1. *Although a reconnection such as the A_1 to R_3 edge in Fig. 3.3 leaves a reaction labeled R_3 from the original model, this is* not *the same reaction as in the original model, and it might be best in fact to relabel any reactions affected by a reduction operation, perhaps with an "m" next to the subscript, e.g. R_{3m}, to indicate a modified reaction. We adopt this convention only when necessary to distinguish a reaction from the original set from a reaction in a transformed bipartite graph. As a consequence, the rate constant of R_{3m} will likely have to be different from the rate constant of R_3 in order for the reduced model to remain in the same dynamical regime as the original model. If we consider specifically the case of removing some positive paths from a long chain of positive paths, the appropriate rate constant for the reconnection will be determined by the rate-limiting step of the chain. Given that the chain may contain reactions of different orders, this may not be as simple as picking out the smallest rate constant, which would not be comparable quantities if they have different units.*

Corollary 1 (Positive Paths in Linked Cycles)**.** *Assume we have a bipartite graph that contains an integrated linked cycle of sufficiently large order, with C_L being the linked cycle. Let C_{L_u} be the collection of all unidirectional cycles in C_L. Let $C_{L_u}[a]$ be an arbitrary*

cycle from C_{L_u}. Then Theorem 1 can be applied to a positive path in $C_{L_u}[a]$ assuming all conditions in Theorem 1 are fulfilled in addition to maintaining unidirectionality of all cycles in C_{L_u}.

Proof. The proof is essentially the same as for Theorem 1. The key is that the number of subgraphs and the values of the corresponding coefficients in any fragment containing the integrated linked cycle are invariant with respect to the edge removals described by Theorem 1. □

Refer to Fig. 3.2(b) for the interpretation of the symbols in Theorem 2.

Theorem 2 (Consecutive Negative Paths in Unidirectional Isolated Cycles)**.** *Assume we have a bipartite graph that contains an integrated unidirectional isolated cycle C_i of sufficiently large order n. Let $a_1, a_2, \ldots, a_n$ be the weights of all n arrows oriented in the direction of C_i, with W_o the set containing the weights a_{h-1}, a_{h+1}, a_{h+3}, and W_e the set containing the weights a_h, a_{h+2}, a_{h+4}. Let $\overline{[A_k, R_j, A_i]}_{a_h, a_{h+1}}$ and $\overline{[A_i, R_m, A_p]}_{a_{h+2}, a_{h+3}}$ be two consecutive negative paths in C_i, where neither of the reaction vertices nor at least two of the species vertices is an entry vertex. Additionally, let R_p and R_s be the reaction vertices immediately preceding and succeeding, respectively, this pair of consecutive negative paths. We have the following cases:*

1. *Neither A_k nor A_i are entry vertices, and at most one element in W_o is non-unitary. Let $\omega_o = \max(W_o)$. If $a_h = a_{h+2} = 1$, then construct the resulting cycle $\overline{C}_i$ as follows: Remove species vertices A_k and A_i as well as the reaction vertices R_j and R_m along with the consecutive arrows with weights a_{h-1} to a_{h+3}, then connect R_p and A_p with an arrow whose direction is that of the original arrow connecting R_p to A_k and weight equal to ω_o.*

2. *Neither A_k nor A_p are entry vertices, and at most one element in each of the sets W_o and W_e is non-unitary. Let $\omega_o = \max(W_o)$ and $\omega_e = \max(W_e)$. Construct the resulting cycle $\overline{C}_i$ as follows: Remove species vertices A_k and A_p and reaction vertices R_j and*

R_m along with the consecutive arrows with weights a_{h-1} to a_{h+4}, then connect R_p and A_i by an arrow with weight ω_o and in the same direction as the original arrow connecting R_p to A_k, and connect A_i and R_s with an arrow of weight ω_e oriented in the direction of the cycle [clockwise as the cycle is drawn in Fig. 3.2(b)].

3. *Neither A_i nor A_p are entry vertices, and at most one element in W_e is non-unitary. Let $\omega_e = \max(W_e)$. If $a_{h+1} = a_{h+3} = 1$, then construct the resulting cycle $\overline{C}_i$ as follows: Remove species vertices A_i and A_p as well as the reaction vertices R_j and R_m along with the consecutive arrows with weights a_h to a_{h+4}, then connect A_k and R_s by an arrow with weight ω_e oriented in the direction of the cycle [clockwise as the cycle is drawn in Fig. 3.2(b)].*

Then any fragment that originally contained C_i will have the same coefficient after replacing C_i by $\overline{C}_i$.

Proof. Let C_i be the cycle of order n corresponding to the integrated unidirectional isolated cycle in a bipartite graph. Then, by definition of integrated unidirectional isolated cycles and subgraphs, this cycle contains an equal number of species and reaction vertices and will appear in any subgraph either as the set of reactant edges or as the entire cycle. In each of the cases in Theorem 2, all removed vertices are non-entry vertices. Thus, removal of these vertices will not affect connectivity to any vertices outside of C_i.

In case 1, A_k and A_i are non-entry vertices and $a_h = a_{h+2} = 1$. Thus, the reactant edges $[\mathrm{A}_k,\ \mathrm{R}_j]_{a_h}$ and $[\mathrm{A}_i,\ \mathrm{R}_m]_{a_{h+2}}$ contribute factors of unity to the coefficient of any subgraph that includes the edge subgraph of C_i. Furthermore, in any subgraph containing C_i, arrows with weights a_{h-1}, a_h, a_{h+1}, a_{h+2}, and a_{h+3} contribute a magnitude of $(a_{h-1})(a_h)(a_{h+1})(a_{h+2})(a_{h+3}) = \omega_o$ to K_C, and therefore to the coefficient of any subgraph in which C_i appears. Connecting $\mathrm{R_p}$ to A_p with an arrow of weight ω_o clearly preserves the magnitude of K_C. Choosing the direction of the new arrow as that of the original $\mathrm{R_p}$ to A_k arrow preserves the parity of K_C and thus its sign. Accordingly, the coefficients of

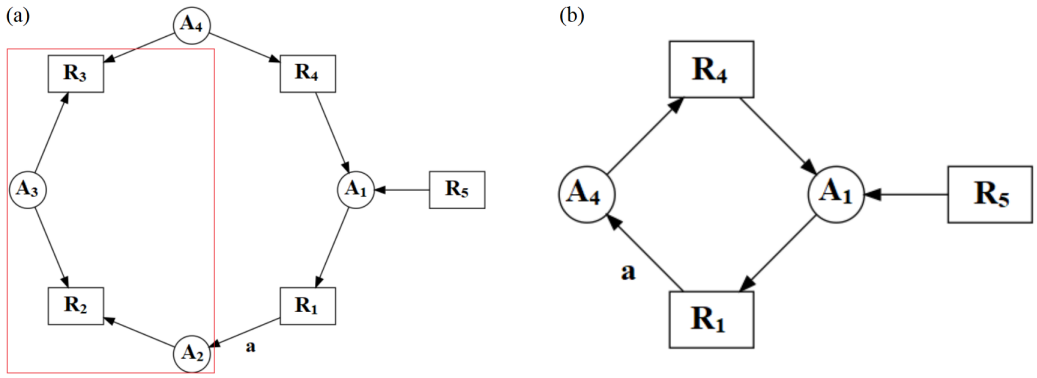

Figure 3.4: (a) An integrated unidirectional isolated cycle that meets the criteria of Theorem 2. For simplicity, we choose only one of the arrows to have an arbitrary weight, $a_{h-1} = a$. The vertices in the red box are the vertices that will be removed. (b) The result of applying Theorem 2 to the figure on the left.

all subgraphs containing C_i will remain unchanged after replacing C_i by $\overline{C}_i$. The number of subgraphs of a fragment containing C_i or its edge subgraph also does not change because of the unidirectionality of the cycle.

Cases 2 and 3 can be proven similarly, the key idea being that both the number and coefficients of subgraphs in a cycle containing C_i do not change post-reduction. □

Figure 3.4 illustrates the use of Theorem 2. It is easily verified that $K_C = a$ for both cycles and their corresponding edge subgraphs both have coefficients of unity.

Corollary 2. *Assume we have a bipartite graph that contains an integrated linked cycle of sufficiently large order, with C_L being the linked cycle. Let C_{L_u} be the collection of all unidirectional cycles in C_L. Let $C_{L_u}[a]$ be an arbitrary cycle from C_{L_u} that does not contain an entry reaction vertex that participates in a negative path with a species vertex in $C_{L_u}[a]$ and a species vertex outside of $C_{L_u}[a]$. Then Theorem 2 can be applied to consecutive negative paths in $C_{L_u}[a]$ assuming all conditions in Theorem 2 are fulfilled in addition to maintaining unidirectionality of all cycles in C_{L_u}.*

Proof. Assuming the extra condition is fulfilled, then the proof of Theorem 2 applies essentially unchanged. That is, all of the cycles in C_L will retain their parities and K_C values

post-reduction. Accordingly, the number of subgraphs in any fragment and their coefficients also remain unchanged. □

Refer to Fig. 3.2(c) for the interpretation of the symbols in Theorem 3.

Theorem 3 (Disjoint Negative Paths in Unidirectional Isolated Cycles)**.** *Assume we have a bipartite graph that contains an integrated unidirectional isolated cycle C_i of sufficiently large order n. Let $a_1, a_2, \ldots, a_n$ be the weights of all n arrows oriented in the direction of C_i. Let $\overline{[A_k, R_j, A_i]}_{a_h, a_{h+1}}$ and $\overline{[A_p, R_m, A_q]}_{a_f, a_{f+1}}$, be two disjoint negative paths in this isolated cycle where neither of the reaction vertices nor at least one species vertex from each negative path is an entry vertex. Denote by W_{o_h} the set containing the weights a_{h-1} and a_{h+1}, by W_{e_h} the set containing the weights a_h and a_{h+2}, by W_{o_f} the set containing the weights a_{f-1} and a_{f+1}, and by W_{e_f} the set containing the weights a_f and a_{f+2}. Additionally, let R_{p_h}, R_{s_h} be the reaction vertices immediately preceding and succeeding, respectively, $\overline{[A_k, R_j, A_i]}_{a_h, a_{h+1}}$, and R_{p_f}, R_{s_f} be the reaction vertices immediately preceding and succeeding, respectively, $\overline{[A_p, R_m, A_q]}_{a_f, a_{f+1}}$.*[8] *Consider the following cases:*

1. *Neither A_k nor A_p are entry vertices and at most one element in each of W_{o_h} and W_{o_f} is non-unitary. Let $\omega_{o_h} = \max(W_{o_h})$ and $\omega_{o_f} = \max(W_{o_f})$. If a_h and a_f are unitary, then construct the resulting cycle $\overline{C}_i$ as follows: Remove species vertices A_k and A_p as well as the reaction vertices R_j and R_m along with the arrows with weights a_{h-1}, a_h, a_{h+1}, a_{f-1}, a_f, a_{f+1}, then connect R_{p_h} and A_i by an arrow with weight ω_{o_h} and in the same direction as the original arrow connecting R_{p_h} to A_k, and connect R_{p_f} and A_q with an arrow of weight ω_{o_f} and in the same direction as the original arrow connecting R_{p_f} and A_p.*

2. *Neither A_k nor A_q are entry vertices and at most one element in each of W_{o_h} and W_{e_f} is non-unitary. Let $\omega_{o_h} = \max(W_{o_h})$ and $\omega_{e_f} = \max(W_{e_f})$. If a_h and a_{f+1} are unitary, then construct the resulting cycle $\overline{C}_i$ as follows: Remove species vertices A_k and A_q*

[8] Note that it is possible to have $R_{p_h} \equiv R_{s_f}$ and/or $R_{s_h} \equiv R_{p_f}$. In either scenario, this theorem holds.

as well as the reaction vertices R_j and R_m along with the arrows with weights a_{h-1}, a_h, a_{h+1}, a_f, a_{f+1}, a_{f+2}, then connect R_{p_h} and A_i by an arrow with weight ω_{o_h} and in the same direction as the original arrow connecting R_{p_h} to A_k, and connect A_p and R_{s_f} with an arrow of weight ω_{e_f} oriented in the direction of the cycle [clockwise as the cycle is drawn in Fig. 3.2(c)].

3. *Neither A_i nor A_p are entry vertices and at most one element in each of W_{e_h} and W_{o_f} is non-unitary. Let $\omega_{e_h} = \max(W_{e_h})$ and $\omega_{o_f} = \max(W_{o_f})$. If a_{h+1} and a_f are unitary, then construct the resulting cycle $\overline{C}_i$ as follows: Remove species vertices A_i and A_p as well as the reaction vertices R_j and R_m along with the arrows with weights a_h, a_{h+1}, a_{h+2}, a_{f-1}, a_f, a_{f+1}, then connect A_k and R_{s_h} by an arrow with weight ω_{e_h} oriented in the direction of the cycle [clockwise as the cycle is drawn in Fig. 3.2(c)], and connect R_{p_f} and A_q with an arrow of weight ω_{o_f} and in the same direction as the original arrow connecting R_{p_f} and A_p.*

4. *Neither A_i nor A_q are entry vertices and at most one element in each of W_{e_h} and W_{e_f} is non-unitary. Let $\omega_{e_h} = \max(W_{e_h})$ and $\omega_{e_f} = \max(W_{e_f})$. If a_{h+1} and a_{f+1} are unitary, then construct the resulting cycle $\overline{C}_i$ as follows: Remove species vertices A_i and A_q as well as the reaction vertices R_j and R_m along with the arrows with weights a_h, a_{h+1}, a_{h+2}, a_f, a_{f+1}, a_{f+2}, then connect A_k and R_{s_h} by an arrow with weight ω_{e_h}, and connect A_p and R_{s_f} by an arrow with weight ω_{e_f}, both oriented in the direction of the cycle [clockwise as the cycle is drawn in Fig. 3.2(c)].*

Then any fragment that originally contained C_i will have the same coefficient after replacing C_i by $\overline{C}_i$.

Proof. By definition of integrated unidirectional isolated cycles, C_i contains an equal number of species and reaction vertices and will appear in subgraphs of any fragment that includes it either as the set of its reactant edges or as the entire cycle. In each case considered above, all removed vertices are non-entry vertices; thus, removal of these will not affect

connectivity to the rest of the vertices outside of C_i. By assumption, we have that R_j and R_m are non-entry vertices.

In case 1, we have the non-entry vertices A_k and A_p and $a_h = a_f = 1$. The reactant edges $[\mathrm{A}_k, \mathrm{R}_j]_{a_h}$ and $[\mathrm{A}_p, \mathrm{R}_m]_{a_f}$ thus contribute factors of unity when the cycle's reactant edge set appears in a subgraph. Furthermore, in any subgraph containing the entire cycle, arrows with weights a_{h-1}, a_h, a_{h+1}, a_{f-1}, a_f, and a_{f+1} contribute a magnitude of $\omega_{o_h}\omega_{o_f}$ to the coefficient of the subgraph. Applying these weights to, respectively, the new product edges $[\mathrm{R}_{\mathrm{p}_h}, \mathrm{A}_i]$ and $[\mathrm{R}_{\mathrm{p}_f}, \mathrm{A}_q]$ means that the coefficient of a subgraph containing the simplified cycle $\overline{C}_i$ will have a coefficient of the same magnitude as the original subgraph. A pair of negative paths is removed, and the rule for the directions of the arrows reconnecting the graph ensures that there is no other change in the number of negative paths, i.e. C_i and $\overline{C}_i$ must have the same parity. Thus, the coefficient of a subgraph containing the entire cycle is preserved. Moreover, the number of subgraphs of a fragment containing C_i or its edge subgraph also does not change after simplification because of the unidirectionality of both C_i and $\overline{C}_i$, which implies that the coefficient of any fragment containing the cycle is preserved.

Cases 2, 3, and 4 can be proven by analogous arguments and are not presented for brevity. □

See Fig. 3.5 for an example of utilizing this theorem.

Corollary 3. *Assume we have a bipartite graph that contains an integrated linked cycle of sufficiently large order, with C_L being the linked cycle. Let C_{L_u} be the collection of all unidirectional cycles in C_L. Let $C_{L_u}[a]$ be an arbitrary cycle from C_{L_u} that does not contain an entry reaction vertex that participates in a negative path with a species vertex in $C_{L_u}[a]$ and a species vertex outside of $C_{L_u}[a]$. Then Theorem 3 can be applied to disjoint negative paths in $C_{L_u}[a]$ assuming all conditions in Theorem 3 are fulfilled in addition to maintaining unidirectionality of all cycles in C_{L_u}.*

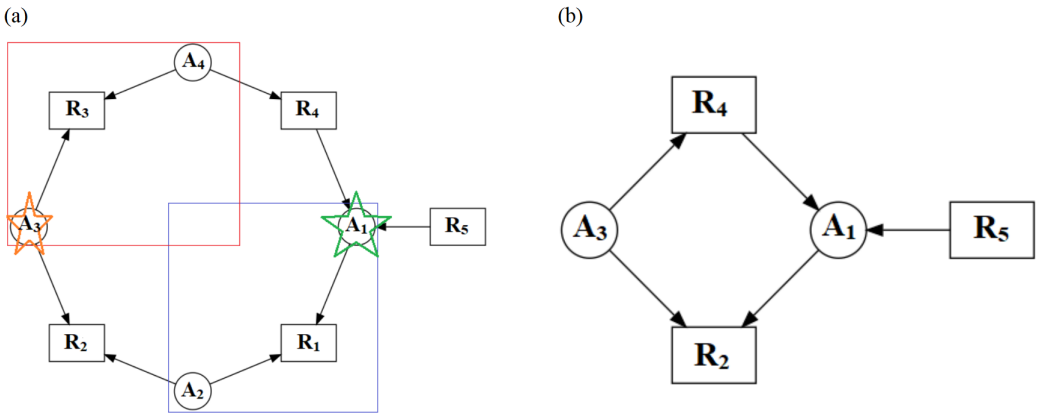

Figure 3.5: (a) An integrated unitary unidirectional isolated cycle that meets the criteria of Theorem 3. The vertices in the red and blue boxes are the disjoint negative paths $\overline{[A_1, R_1, A_2]}$ and $\overline{[A_3, R_3, A_4]}$. The species vertex with the green star (A_1) will be kept from the blue negative path and the species vertex with the orange star (A_3) will be kept from the red negative path. (b) The result of applying Theorem 3 to the figure on the left.

Proof. Assuming the extra condition is fulfilled, then the same proof for Theorem 3 applies. That is, all of the cycles in C_L will still be present post-reduction, their contributions to the coefficients of various subgraphs will be the same, and the number of subgraphs in a fragment containing these cycles will be unchanged. □

3.3.2 (Bio)chemical Relevance

The theorems and corollaries found in the previous section describe reductions that preserve critical fragments given a bipartite graph that meets a theorem's conditions. They do not, however, address the (bio)chemical relevance of the reduced model. For example, it is possible that a removed species or reaction vertex may have been a well-studied component of a mechanism and that there are some observations regarding this component that a model should be able to match. In these cases, it may be desirable to identify species to be preserved prior to any reductions being made, and to avoid eliminating these species in the course of reduction. In many cases, especially if most of the edges are unitary, the theorems presented above offer multiple options for reduction. It will accordingly sometimes be possible to carry out reductions while protecting specific species or reactions.

Unfortunately, reductions involving negative paths are not as straightforward as reduc-

tions of positive paths because a negative path encodes a co-reactant relationship, the product not being included in an isolated cycle. In this case, the product species (if any) would *have* to be outside the cycle. Theorems 2 and 3 remove *both* reaction vertices involved in the negative paths subject to reduction. Unless the product is formed by another reduction, this therefore eliminates the formation of the product from the model. If the product is dynamically important, then its removal may sabotage the model altogether. However, in order for the product to play an important dynamical role, it would have to be part of a cycle, and thus the cycle on which we were operating would not be isolated, and the reaction vertex forming the product would in fact be a non-trivial entry vertex, which our theorems explicitly disallow as targets for reduction. The circumstances in which it would be permissible to remove a vertex producing a product would involve products that are, in the language of chemical reaction network theory, part of a terminal linkage class [6], i.e. a pathway that acts as a sink. This includes the trivial case in which the immediate product is a sink species, often represented by an empty-set symbol in reaction mechanisms.

There is also the significant problem of interpreting the bipartite graph resulting from the elimination of negative paths. For instance, consider the example displaying the use of Theorem 3 shown in Fig. 3.5. The original cycle encoded the reactions (neglecting species not explicitly shown in the cycle) $A_1 + A_2 \rightarrow$, $A_2 + A_3 \rightarrow$, $A_3 + A_4 \rightarrow$ and $A_4 \rightarrow A_1$. In the reduced cycle on the other hand, we have $A_1 + A_3 \rightarrow$ and $A_3 \rightarrow A_1$. These are quite different in terms of their chemical meanings, and it may be that, for example, there is no chemical sense to be attached to A_1 reacting with A_3. Even if that is not the case, the reactant-product relationships have been profoundly transformed and, arguably, the reduced model is not a model for the original reaction but for a different reaction. An alternative perspective is that we do not care about the chemical interpretation of the reduced bipartite graph, only that it is easier to analyze. The discovery of a critical fragment in the reduced model could then be mapped back to the original model.

As an interesting side-note, in the classical theory of kinetic equivalence of mechanisms,

two models are kinetically equivalent if they are related by a linear transformation [27–29]. The reactions in the reduced model will always be linear combinations of reactions in the original model provided we consider *only* the reactions in the cycle. The application of Theorem 1 essentially involves "squeezing out" intermediates in a cycle, which can be accomplished by taking a linear superposition of reactions with positive coefficients. The situation with negative paths is more complicated, but this is essentially equivalent to subtracting reactions. For the example of Fig. 3.5, if we call the resulting reactions R_{2m} and R_{4m}, we have $R_{2m} = R_3 - R_4$ and $R_{4m} = R_2 - R_1$.

3.4 Application to a Gene Expression Control Model

3.4.1 Model Description

A delayed mass-action model for the transcriptional control of Hmp, an NO detoxifying enzyme, by the iron-sulfur protein FNR displays bistability [30]. FNR acts as a transcriptional repressor, preventing the synthesis of the *hmp* mRNA when bound to the gene's promoter. Nitric oxide (NO), when present in the cell, reacts with the iron-sulfur cluster of FNR. This causes FNR to lose its ability to bind the *hmp* promoter, and thus allows this gene to be transcribed. The synthesized Hmp can then remove NO, converting it to nitrate (NO_3^-), which is much less toxic. A model of this system is complicated by the potential for up to 8 molecules of NO to react with a single iron-sulfur cluster [31]. Only four of these reactions are kinetically distinguishable, leading to a model with five nitrosylation states of FNR [30]. The original model contains 18 species and 32 reactions. Both expression and clearance delays were considered, the latter representing the time during which the promoter is occluded by an RNA polymerase or the ribosome binding site on an mRNA molecule is occluded by a ribosome. Clearance delays turn out to have significant dynamical effects [32]. Nevertheless, in the present contribution, an ODE model was considered given that the *potential* for bistability does not depend on the inclusion of delays, thus providing an interesting example to which to apply the theorems presented

here. Beyond the neglect of delays, we also considered a simplified version of the original model for two reasons: Perhaps most importantly, an example based on a large model with a complex bipartite graph is poorly suited as an illustration. The original model is also extremely densely connected, largely through NO. As a result, paradoxically, our model simplification methods apply to very few species and reactions in the pre-existing model. It should be noted that this pattern of dense connections is unusual. Most species in metabolic networks, for example, have very few connections to other species via reactions [33]. The simplified model presented below is thus, we feel, more representative of models likely to be encountered in practice.

The simplified mechanism consists of the following 15 mass-action reactions connecting 10 species, not counting the sink species NO_3^-:

$$\xrightarrow{k_{\text{in}}} \text{NO} \tag{3.1a}$$

$$\text{Hmp} \xrightarrow{k_1} \text{Hmp}\cdot\text{O}_2 \tag{3.1b}$$

$$\text{Hmp}\cdot\text{O}_2 + \text{NO} \xrightarrow{k_2} \text{Hmp}\cdot\text{O}_2\cdot\text{NO} \tag{3.1c}$$

$$\text{Hmp}\cdot\text{O}_2\cdot\text{NO} \xrightarrow{k_{-3}} \text{Hmp} + \text{NO}_3^- \tag{3.1d}$$

$$\text{Hmp} + \text{NO} \xrightarrow{k_4} \text{Hmp}\cdot\text{NO} \tag{3.1e}$$

$$\text{Pro}^{hmp} \xrightarrow{k_{\text{tc}}} \text{Pro}^{hmp} + \text{RBS}^{hmp} \tag{3.1f}$$

$$\text{RBS}^{hmp} \xrightarrow{k_{\text{tl}}} \text{RBS}^{hmp} + \text{Hmp} \tag{3.1g}$$

$$\text{RBS}^{hmp} \xrightarrow{k_5} \emptyset \tag{3.1h}$$

$$\text{Hmp} \xrightarrow{k_{6a}} \emptyset \tag{3.1i}$$

$$\text{Hmp}\cdot\text{O}_2 \xrightarrow{k_{6b}} \emptyset \tag{3.1j}$$

$$\text{Hmp}\cdot\text{NO} \xrightarrow{k_{6d}} \text{NO} \tag{3.1k}$$

$$\text{Pro}^{hmp} + \text{FNR} \underset{k_{-A}}{\overset{k_A}{\rightleftharpoons}} \text{Pro}^{hmp}\cdot\text{FNR} \tag{3.1l}$$

$$\text{FNR} + \text{NO} \xrightarrow{k_7} \text{FNR}_e \tag{3.1m}$$

$$\mathrm{FNR}_e \xrightarrow{k_r} \mathrm{FNR} \tag{3.1n}$$

Reaction (3.1a) represents a source of nitric oxide, which may be endogeneous or exogeneous, i.e. due to diffusion of NO into the cell from the extracellular medium. Reactions (3.1b) to (3.1e) describe the kinetics of the conversion of NO to nitrate (NO_3^-) catalyzed by Hmp. In reaction (3.1b), a molecule of oxygen binds into the active site of the enzyme. Because we hold the concentration of oxygen constant, (3.1b) is represented as a pseudo-first-order reaction with effective rate constant $k_1 = k_{O_2}[O_2]$, where k_{O_2} is the true second-order rate constant for this binding step. Reaction (3.1f) represents the transcription of the *hmp* gene, starting at its promoter (Pro^{hmp}), generating an mRNA. We focus particularly on the ribosome binding site (RBS) because translation [reaction (3.1g)] can be initiated as soon as the RBS becomes available in bacteria. Moreover, mRNAs are typically degraded starting from the RBS in bacteria [34] [reaction (3.1h)]. Reactions (3.1i) to (3.1k) are enzyme degradation reactions. FNR is a repressor of *hmp* expression, as shown in reaction (3.1l). Reaction (3.1m) depicts the inactivation of FNR by nitric oxide, and reaction (3.1n) represents a recycling pathway. Details of the biochemistry are given in Ref. [30].

The model presented above involves the following simplifications from the original:

- $[O_2]$ was a parameter in the original model. The value of this concentration was therefore combined with the original k_1 to define a new pseudo-first-order rate constant, which appears in reaction (3.1b) of the current model. We retain the label k_1 for this new rate constant.

- The reactions leading to the detoxification of NO by Hmp, reactions (3.1b) to (3.1d), were assumed to be irreversible, as was the formation of the inhibitory complex $\mathrm{Hmp}\cdot\mathrm{NO}$ in reaction (3.1e), which now appears in the mechanism as a suicide substrate inhibition reaction [35].

- We assume that the $\mathrm{Hmp}\cdot O_2\cdot\mathrm{NO}$ complex formed in reaction (3.1c) is stable, i.e. we

do not have a sink for this species. This simply reduces the number of reactions to be considered by one without having any significant dynamical effect.

- The multistep degradation of the iron-sulfur cluster of FNR has been reduced to the single reaction (3.1m).

Model (3.1) has two conservation relations, one for the gene promoter, and one for FNR. Accordingly, given the 10 chemical species involved in the dynamics, the model has a rank of 8.

The parameters of Table 3.1 were used throughout this study unless otherwise noted.

3.4.2 Analysis of the Model

Fig. 3.6 shows the bifurcation diagram of the model varying k_{in}. Unlike the original model [30], the upper branch of steady states diverges near $k_{\text{in}} = 0.248\,\mu\text{M}\,\text{s}^{-1}$ because of the suicide substrate inhibition replacing the reversible competitive substrate inhibition in the original model. Between $k_{\text{in}} = 0.248$ and the saddle-node bifurcation at $0.375\,\mu\text{M}\,\text{s}^{-1}$, there is only one stable steady state, the lower one, but that steady state has a relatively small basin of attraction. Many initial conditions result in runaway trajectories in which [NO] increases without bound. For values of k_{in} greater than the saddle-node value, only runaway trajectories are observed. There is a second saddle-node point at $k_{\text{in}} = 2.2 \times 10^{-3}\,\mu\text{M}\,\text{s}^{-1}$. Below this value of k_{in}, only the lower steady state exists, and it attracts all trajectories. Between $k_{\text{in}} = 2.2 \times 10^{-3}$ and $0.248\,\mu\text{M}\,\text{s}^{-1}$, the system is bistable.

The bipartite graph of the model is illustrated in Fig. 3.7. This model has five critical fragments of order 8, one of which is shown in Fig. 3.8. The other four critical fragments are not drawn, but only differ by free-floating edges. For example, in one of the other critical fragments, the edge [Pro^{hmp}, R_A] is replaced by [FNR, R_A]. The subgraphs of the critical fragment of Fig. 3.8 are drawn in Fig. 3.9. From these subgraphs, we can calculate a fragment coefficient $K_{S_8} = -1$.

Table 3.1: Parameter values for the original Hmp model and its reduced versions. Adapted from Ref. [30].

Parameter	Value
Hmp Catalysis	
k_1	$128\ \mathrm{s}^{-1}$
k_2	$2.5 \times 10^3\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$
k_{-3}	$6.2 \times 10^2\ \mathrm{s}^{-1}$
k_4	$8.0 \times 10^1\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$
Hmp expression	
$[\mathrm{Pro}^{hmp}]_{\mathrm{total}}$	$1.0 \times 10^{-3}\ \mu\mathrm{M}$
k_{tc}	$0.71\ \mathrm{s}^{-1}$
k_{tl}	$0.83\ \mathrm{s}^{-1}$
k_5	$0.014\ \mathrm{s}^{-1}$
$k_{6a} = k_{6b} = k_{6d}$	$2.4 \times 10^{-4}\ \mathrm{s}^{-1}$
FNR dynamics	
$[\mathrm{FNR}]_{\mathrm{total}}$	$2.5\ \mu\mathrm{M}$
k_7	$1.6 \times 10^{-3}\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$
k_A	$1.0 \times 10^1\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$
k_{-A}	$0.11\ \mathrm{s}^{-1}$
k_r	$0.1\ \mathrm{s}^{-1}$
Parameters in reduced models	
$k_{2\mathrm{m}}$	$2.5 \times 10^3\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$
$k_{4\mathrm{m}}$	$8.0 \times 10^1\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$
$k_{7\mathrm{m}}$	$1.6 \times 10^{-3}\ \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$
k_{tcm}	$5.0 \times 10^1\ \mathrm{s}^{-1}$

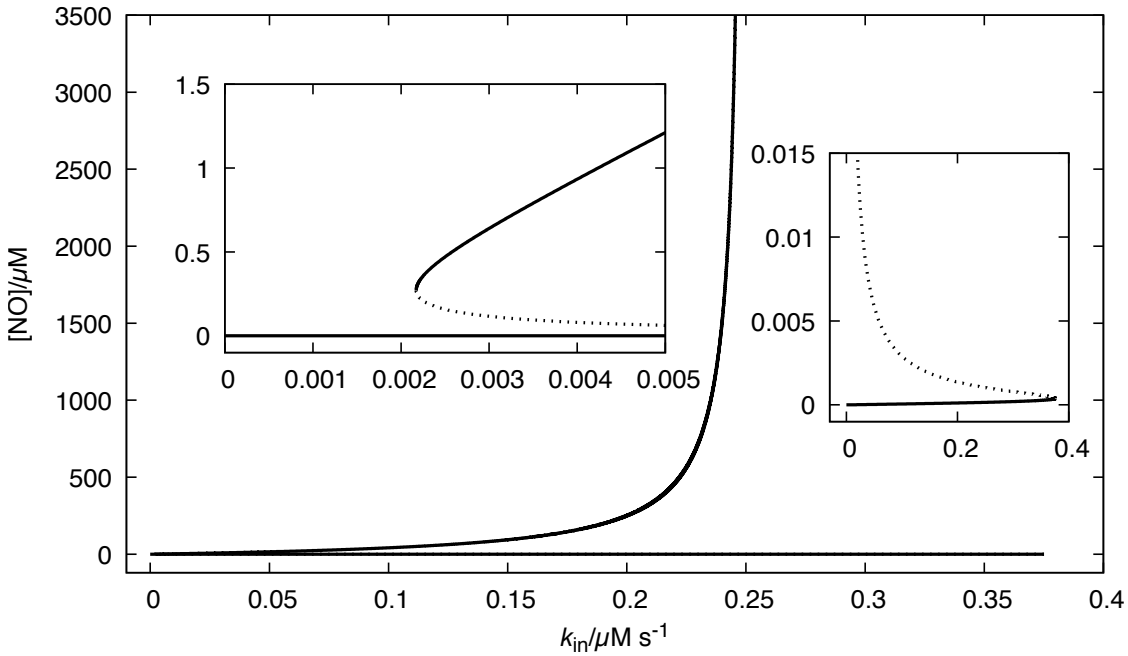

Figure 3.6: Bifurcation diagram for the Hmp model (3.1). In this and subsequent bifurcation diagrams, solid lines represent stable steady states while dotted lines represent unstable steady states. The insets magnify the regions in which the two saddle-node bifurcations occur. Note: all bifurcation diagrams were computed using Xppaut version 8 [36].

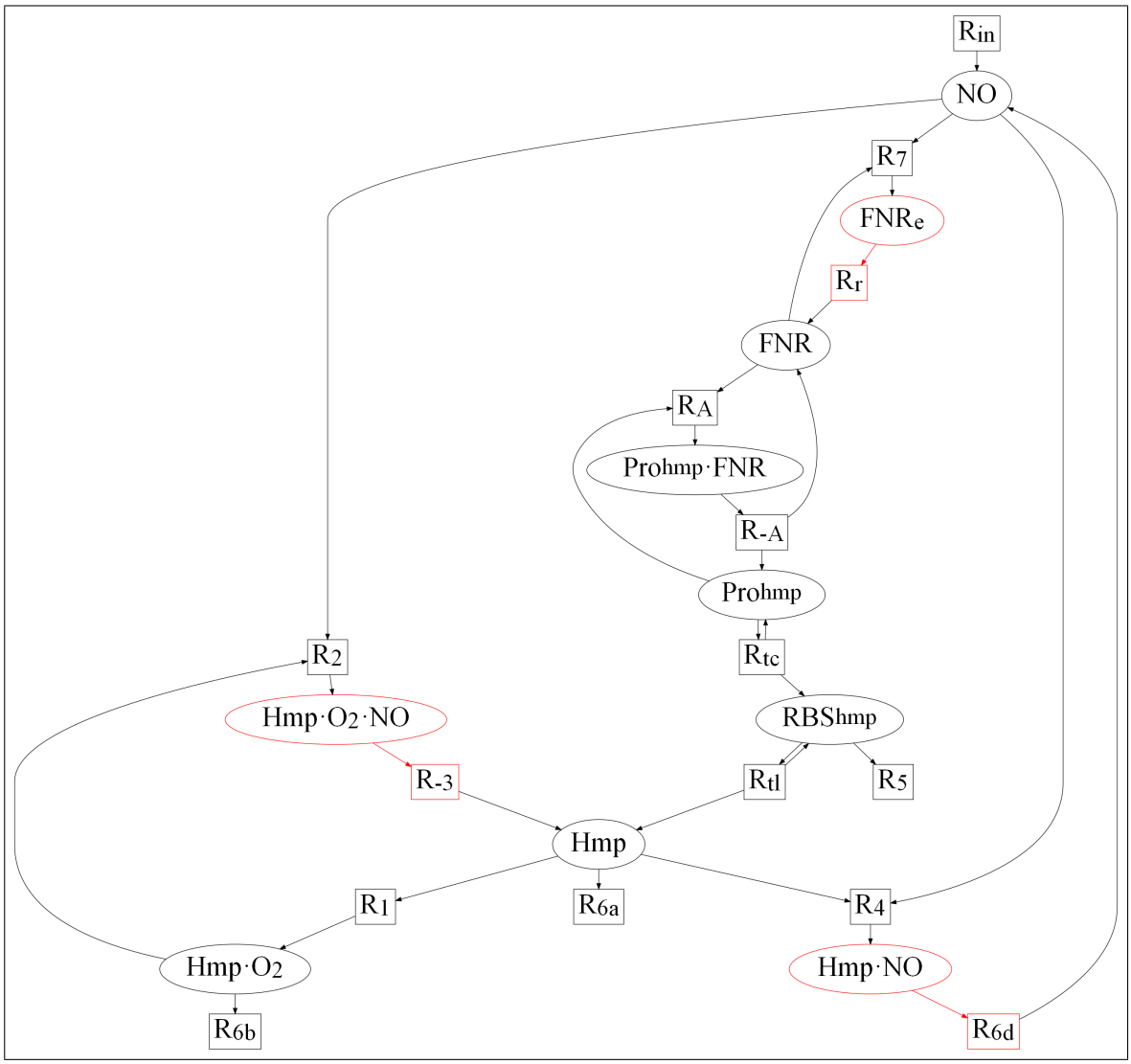

Figure 3.7: The bipartite graph for the Hmp model (3.1). The red edges feature in the reductions of Sect. 3.4.3.

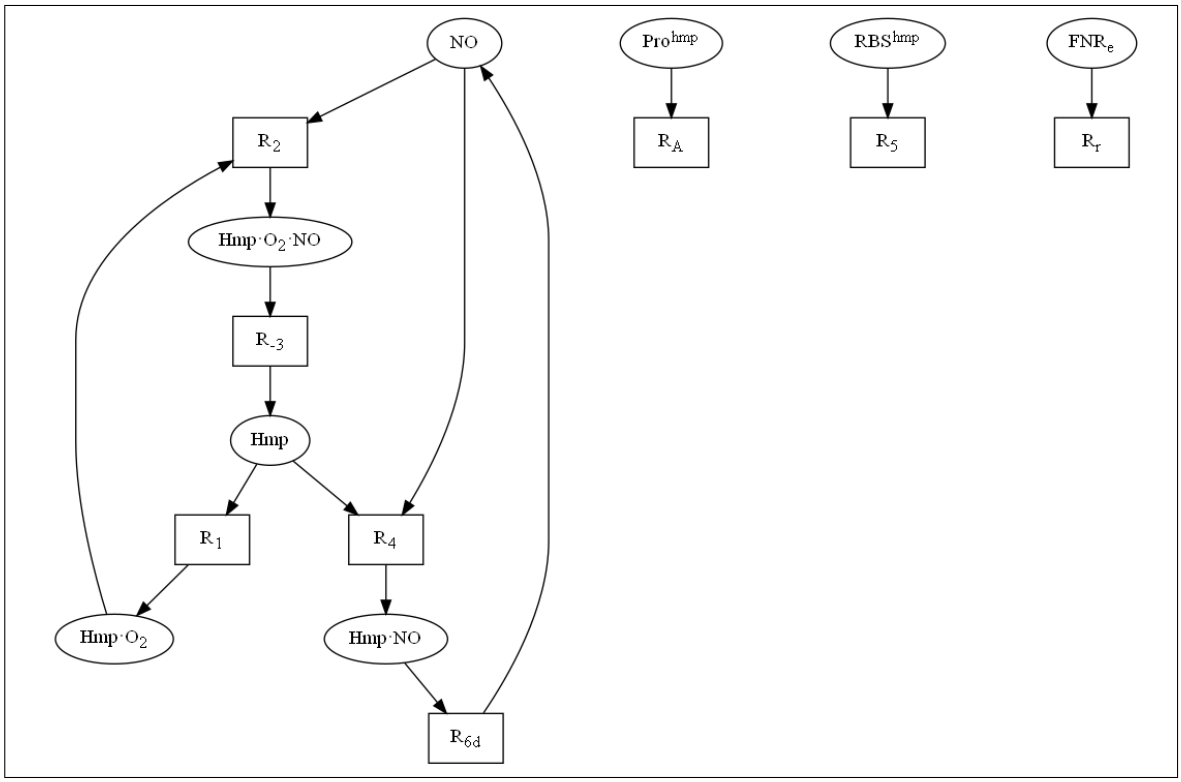

Figure 3.8: One of the five critical fragments of order 8 (full rank) of the Hmp model.

As an added note, consider the structure of the critical fragment in Fig. 3.8, which includes the edges [Pro^{hmp}, R_A] and [RBS^{hmp}, R_5], and [FNR_e, R_7]. Because these edges just contribute multiplicative factors of 1 to the coefficient of the fragment, a critical fragment of order 5 is obtained by leaving these out. In turn, a critical fragment of order less than the rank implies the possibility of an Andronov-Hopf bifurcation [14]. We have not attempted to discover an oscillatory regime in this model. We point this out because the build-up of larger critical fragments from smaller critical fragments is a common theme in the analysis of bipartite graphs, explaining (in part) why it is not uncommon to find both Andronov-Hopf and saddle-node bifurcations in the same model.

3.4.3 Conservative Reductions

The bipartite graph of the model (Fig. 3.7) includes a few linked cycles containing a total of three positive paths to which we can apply Corollary 1. These positive paths

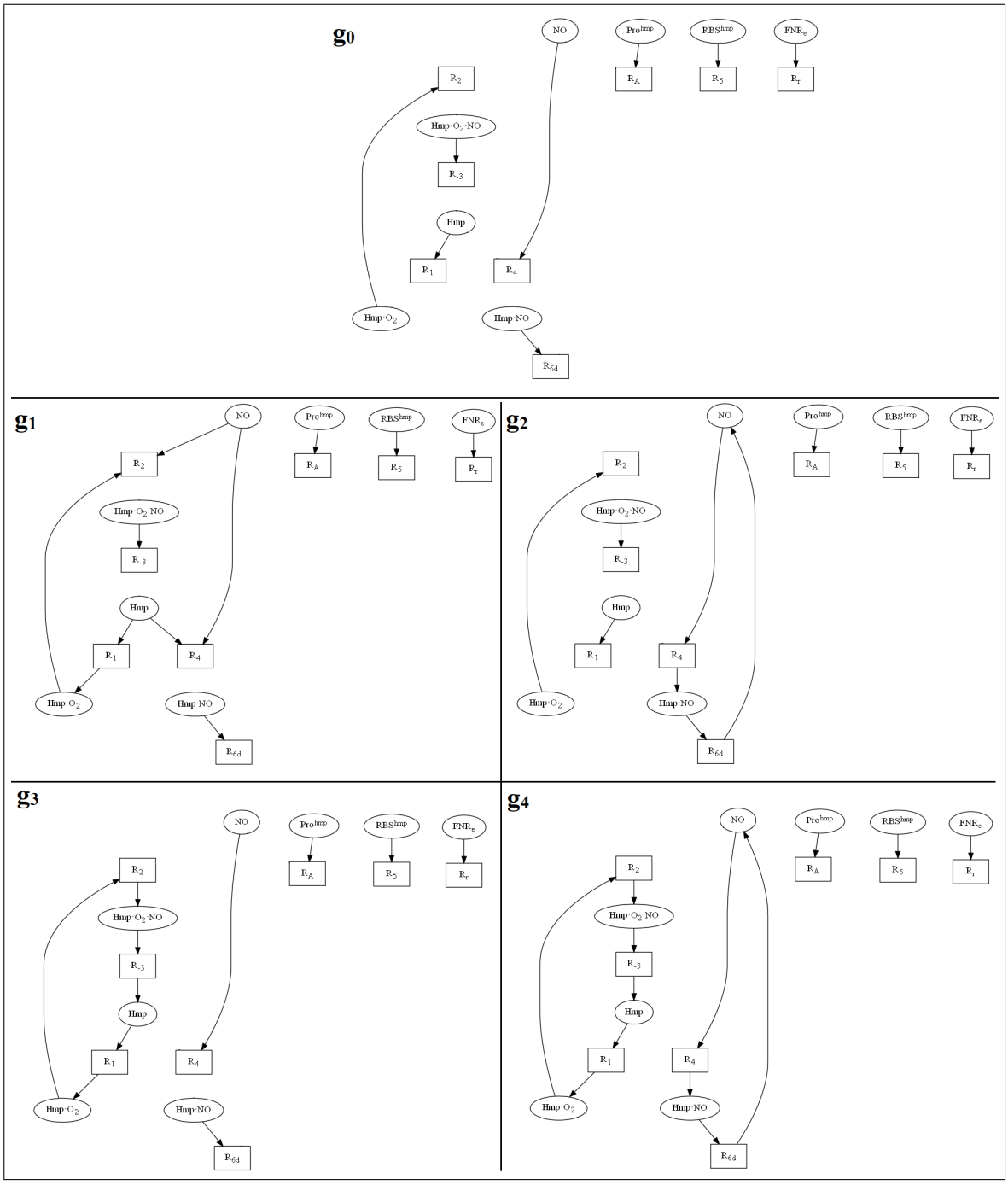

Figure 3.9: The subgraphs of the critical fragment (from Fig. 3.8) of the Hmp model.

are [FNR_e, R_r, FNR], [$Hmp \cdot O_2 \cdot NO$, R_{-3}, Hmp], and [$Hmp \cdot NO$, R_{6d}, NO]. We can remove the reactant edge of each path (indicated in red in the bipartite graph) by making the appropriate connection from the preceding reaction vertex to the succeeding species vertex of each edge. We rename R_2 to R_{2m}, R_4 to R_{4m}, and R_7 to R_{7m} to indicate that the *modified* reaction obtained after each reconnection is different from the original reaction. These connections preserve all of the cycles involved in the linked cycle and change neither the structure nor the coefficients of any of the subgraphs of the system's critical fragments. However, due to the removal of FNR_e and R_r, we lose two critical fragments that included this edge as a free-floating edge. We can think of this kind of loss of critical fragments as the lifting of a degeneracy due to the removal of inessential components that do not participate in the cyclical structures that give rise to criticality. The resulting bipartite graph is shown in Fig. 3.10. Chemically, the effect of these transformations is to replace reactions: (3.1c) and (3.1d); (3.1e) and (3.1k); and (3.1m) and (3.1n) by their respective sums:

$$Hmp \cdot O_2 + NO \xrightarrow{k_{2m}} Hmp + NO_3^-, \tag{3.2a}$$

$$Hmp + NO \xrightarrow{k_{4m}} NO, \tag{3.2b}$$

$$FNR + NO \xrightarrow{k_{7m}} FNR, \tag{3.2c}$$

"squeezing out" the intermediates $Hmp \cdot O_2 \cdot NO$, $Hmp \cdot NO$ and FNR_e, respectively, which, as has previously been noted, is related to classical notions of kinetic equivalence [27–29].

The critical fragment of the simplified model and its subgraphs are shown in Figs. 3.11 and 3.12, respectively. The elimination of the three species and three reactions has reduced the number of active chemical species to 7 and the number of reactions to 12. There are still two conservation relations, so the rank of the model is now 5. Thus, the critical fragment of order 5 shown in Fig. 3.11 is a full-rank critical fragment corresponding to the critical fragment of the original model shown in Fig. 3.8. Comparing the subgraphs (Figs. 3.9 and 3.12), we see that the reduction has preserved

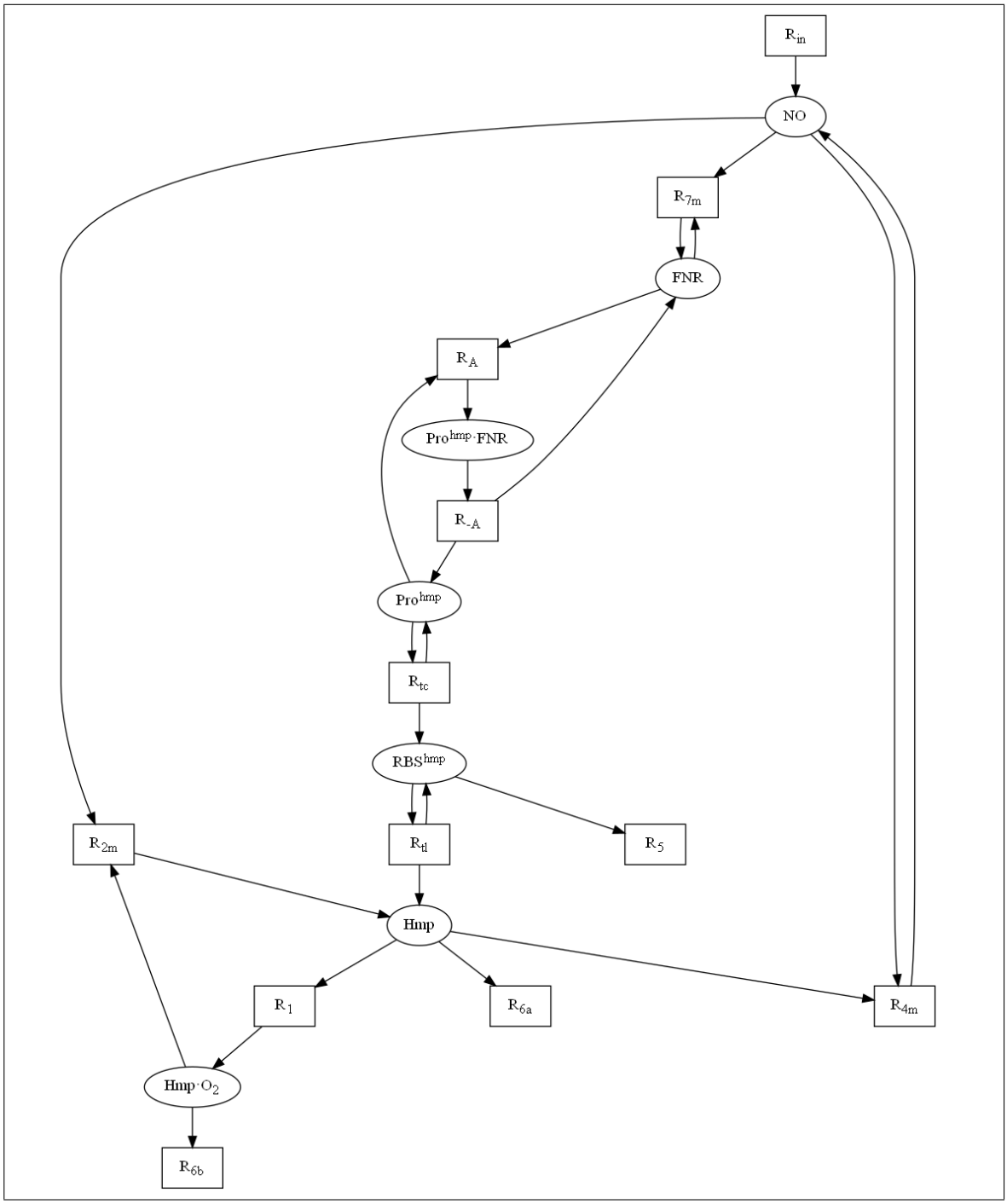

Figure 3.10: Bipartite graph of the model obtained following conservative reductions. This model contains only 7 species (not counting the nitrate sink) and 12 reactions.

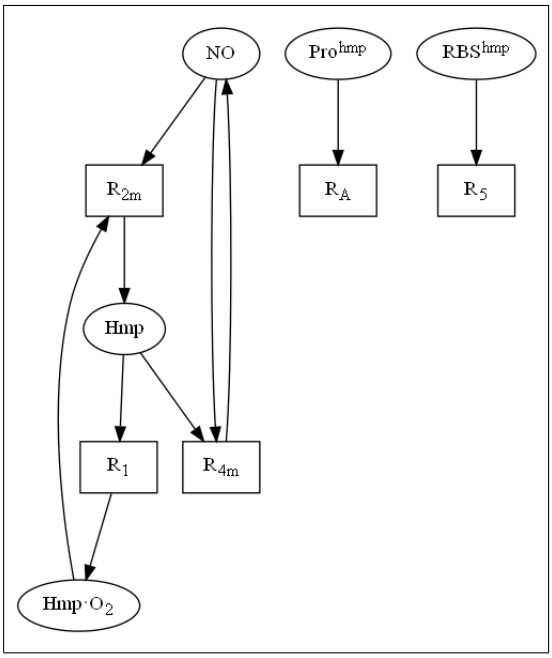

Figure 3.11: One of the three critical fragments of order 5 (full rank) of the Hmp model after conservative reductions. This critical fragment corresponds to the critical fragment of the full model found in Fig. 3.8.

- the number of subgraphs of the critical fragment, and
- the number and parities of cycles in each subgraph.

Thus, the coefficient of each critical fragment in the two models would be the same. Indeed, we can easily calculate $K_{S_5} = -1$ for this fragment. Since we chose a critical fragment for simplification, the fragment shown in Fig. 3.11 is still critical. In fact, the bifurcation diagram of the model retains the two saddle-node bifurcations of the original model in essentially identical positions despite the radical simplifications made (Fig. 3.13; compare Fig. 3.6). Interestingly however, the reduced model has at least one stable steady state for any value of the control parameter k_{in}, unlike the original model in which a large inflow rate of NO can overwhelm the enzyme and its control system. This is doubtless due to the simplification of Michaelis-Menten kinetics to the simple bimolecular reactions (3.2) which lack saturation behavior. Indeed, our graph transformations retain the dynamics of the original model only in the sense that they preserve the capacity for bifurcations, but

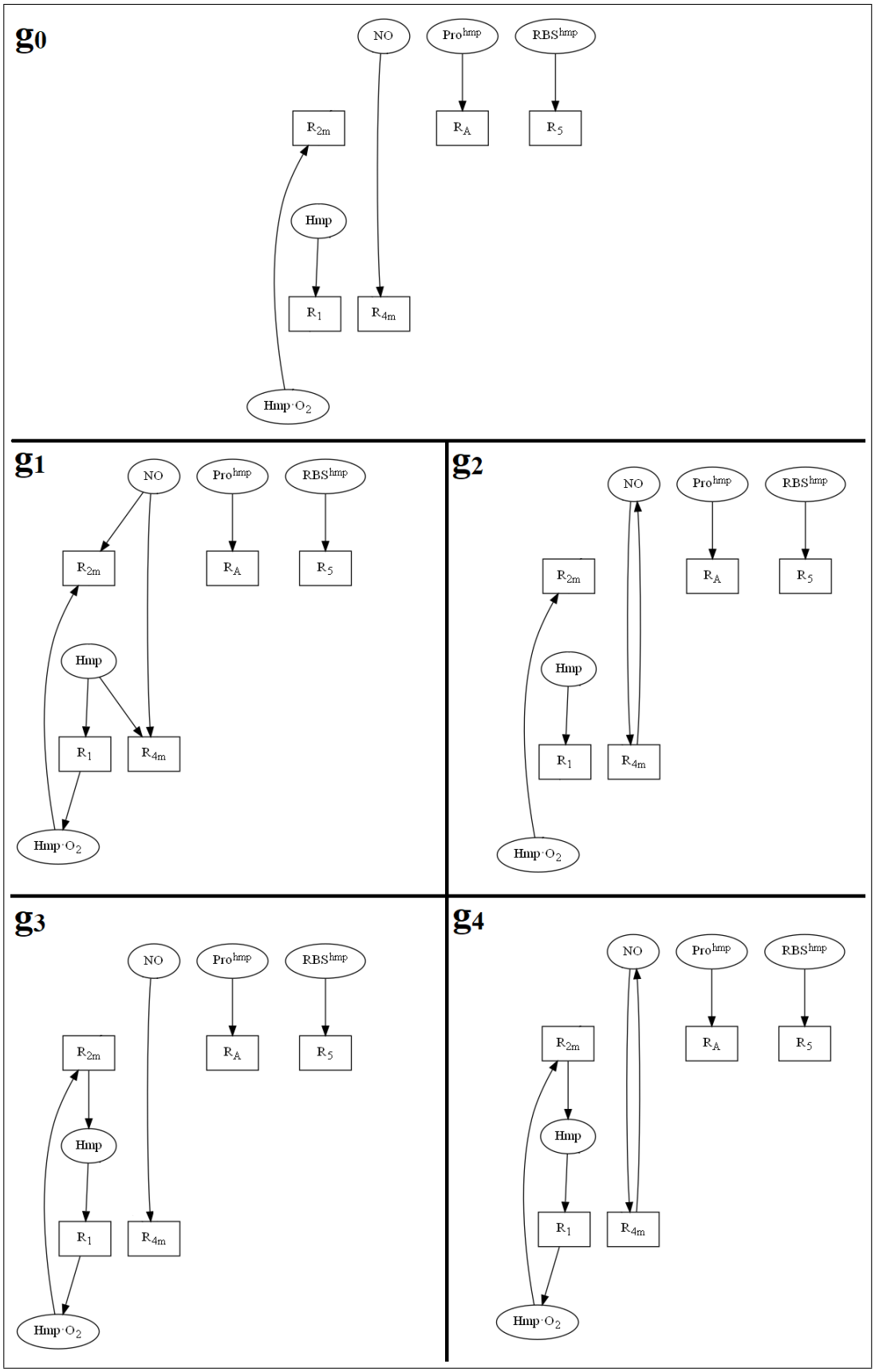

Figure 3.12: The subgraphs of the critical fragment from Fig. 3.11

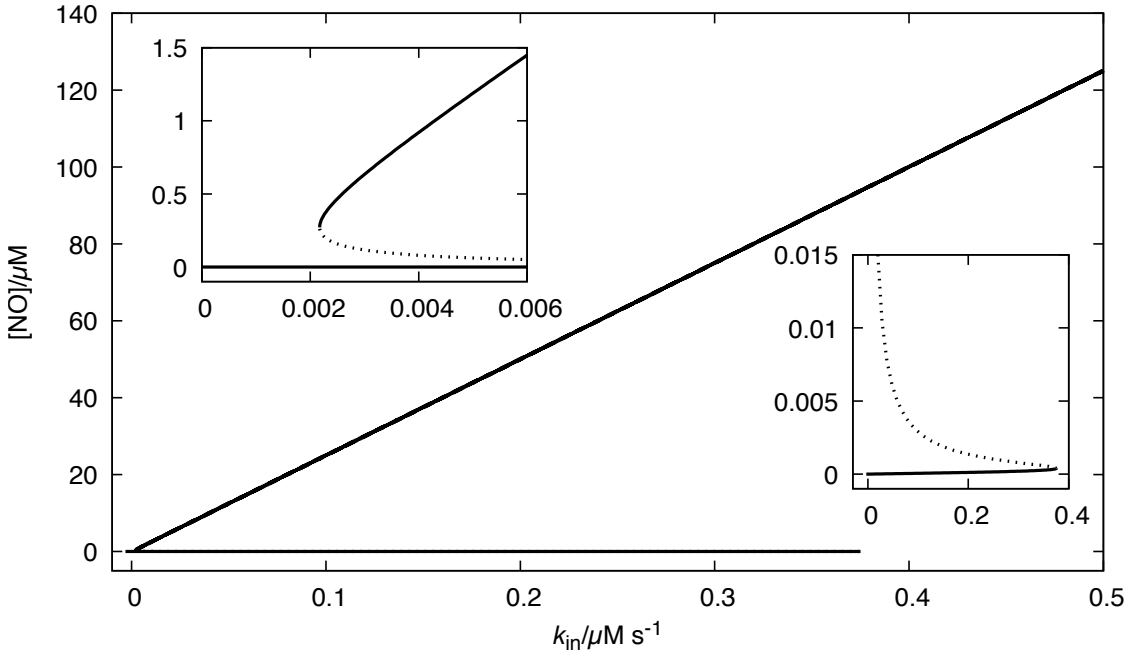

Figure 3.13: Bifurcation diagram for the conservative reduction of the Hmp model shown in Fig. 3.10 using the parameters of Table 3.1 with $k_{2\mathrm{m}} = k_2$, $k_{4\mathrm{m}} = k_4$ and $k_{7\mathrm{m}} = k_7$.

there is no guarantee that the behavior will be identical away from the bifurcation points or for that matter that the bifurcations will occur at similar parameter values, as was found to be the case here.

It is worth pausing to consider the net reaction (3.2c), and to note that this reduction has eliminated the feedback loop between nitrosylation of FNR and control of the synthesis of Hmp. In other words, this control system is not essential to the bistability of this model. In the reduced model, FNR becomes a second enzyme that eliminates NO. As should perhaps have been clear from the original critical fragment (Fig. 3.8), the key interactions leading to bistability are located in the catalytic system, including the substrate inhibition by NO. After reduction, the critical fragment (Fig. 3.11) still contains two competing pathways, one via reactions R_1 and $\mathrm{R}_{2\mathrm{m}}$ that eliminates NO, and the other via reaction $\mathrm{R}_{4\mathrm{m}}$ that eliminates Hmp. An interaction analogous to competitive substrate inhibition has previously been shown to generate oscillations in a model for hydrogen oxidation [37].

3.4.4 Aggressive Reductions

In Sect. 3.4.3, we applied the theorems presented in this thesis to all paths that met the criteria of our conservative reduction theorems and showed that, upon reduction, the resulting model still exhibits bistability. There may however be more reductions available that would preserve bistability based on a knowledge of the critical fragment and a careful examination of the bipartite graph, but these would involve the deletion of entry vertices. We call these *aggressive reductions*. The critical fragment of the reduced model, Fig. 3.11, tells us that we must preserve all of the species in the linked cycles. Anything outside of these linked cycles either doesn't contribute to critical fragments, or contributes only trivially as free-floating edges (e.g. [Pro^{hmp}, R_A]). Provided they are not involved in other significant interactions, we could consider deleting some additional species and reactions even if they do not satisfy the conditions of our theorems. Looking at the bipartite graph (Fig. 3.10), we see for example that RBS^{hmp} only appears as an intermediate in the Hmp synthesis pathway. It is an entry vertex, and it participates in a cycle of order 1 with $\mathrm{R_{tl}}$, so it does not satisfy the conditions for conservative reductions. Nevertheless, knowing that it appears in critical fragments only trivially, we can contemplate its removal. On the other hand, it is not clear how the graph could be reconnected in a sensible way if we removed, say, Pro^{hmp} and $\mathrm{R_{tc}}$, even though the critical fragment suggests these species are not essential to the bistable behavior. A similar comment could be made about FNR and the reactions with which it shares edges. We therefore attempt to remove only RBS^{hmp} and the reactions $\mathrm{R_{tl}}$ and $\mathrm{R_5}$, connecting $\mathrm{R_{tc}}$ directly to Hmp. As per our previously established convention, we rename the modified reaction $\mathrm{R_{tcm}}$. The resulting bipartite graph is shown in Fig. 3.14. From the graph, the modified reaction $\mathrm{R_{tcm}}$ reads as follows:

$$\mathrm{Pro}^{hmp} \xrightarrow{k_{\mathrm{tcm}}} \mathrm{Pro}^{hmp} + \mathrm{Hmp}. \tag{3.3a}$$

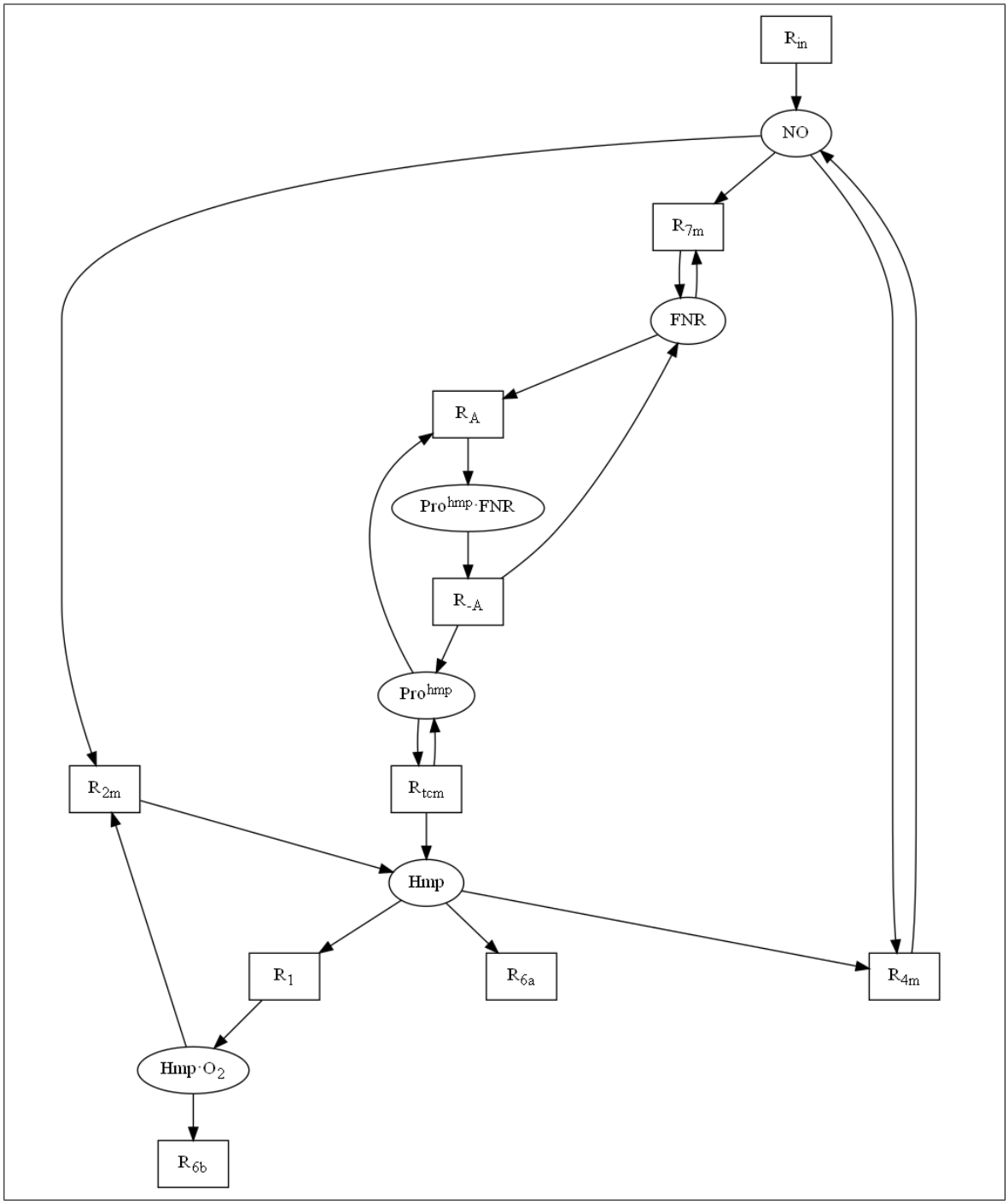

Figure 3.14: The bipartite graph of the aggressively reduced model of Sect. 3.4.4. This model still exhibits bistability but contains only 6 species (not counting the nitrate sink) and 10 reactions. There are still two conservation relations, one for the promoter and one for FNR, so the rank of the model is now 4.

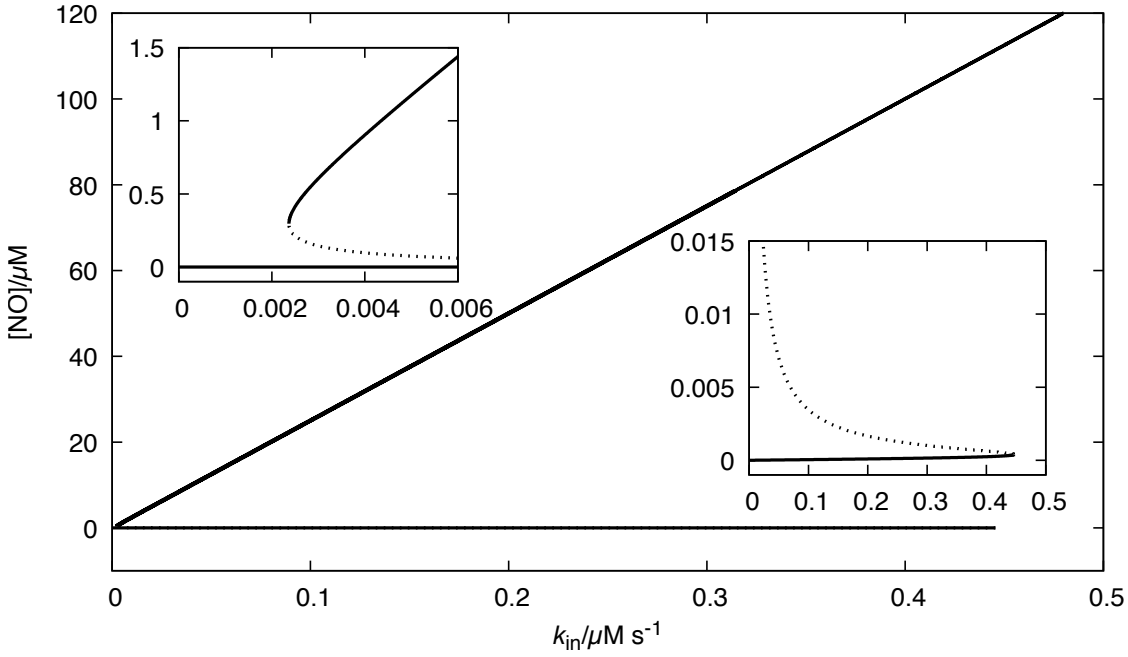

Figure 3.15: Bifurcation diagram of the model following removal of RBS^{hmp} and associated reactions from the model. All parameters are as in Fig. 3.13 except $k_{\mathrm{tcm}} = 50\,\mathrm{s}^{-1}$.

We note in passing that models in which mRNA does not explicitly appear and proteins are shown as being synthesized directly from a gene are not uncommon in the literature. This simplification changes the delay distribution separating gene activation from the appearance of a protein, which could be important in models with Andronov-Hopf bifurcations, but is of little consequence in a system where only saddle-node bifurcations are of interest.

As expected, the critical fragment of this aggressively reduced version of the model is identical to Fig. 3.11, barring the [RBS^{hmp}, R_5] edge. This model still contains three critical fragments but now each of order 4 (full rank). Its bistable regime is depicted in Fig. 3.15. Because mRNA serves (in part) an amplification role (many mRNAs to one promoter), in order to get similar dynamics in the absence of mRNA, we need to increase the transcription initiation rate constant. If we do, we end up with a similar bifurcation diagram to the one found for the model that includes RBS^{hmp}. (Compare Figs. 3.13 and 3.15.)

Up to this point, all model reductions considered have retained a pair of saddle-node bifurcations between which we can observe bistability. We have found that further aggressive

reductions destroy at least one of the saddle-node bifurcations, leaving us with a system that does not have a bistable regime. Note that a critical fragment of full rank is a necessary condition for a saddle-node bifurcation to occur. Because the condition is not sufficient, there is no guarantee that there will be parameters where a saddle-node bifurcation occurs, never mind two, just because the network has a critical fragment.

The [R_1, $Hmp \cdot O_2$] product edge in Fig. 3.14 might seem like an attractive target for an aggressive reduction of a slightly different kind. While this edge does appear in the critical fragment (Fig. 3.11), given that its only other connection is to a sink (R_{6b}), it is tempting to think that we could remove $Hmp \cdot O_2$ along with the reaction vertices R_1 and R_{6b}, closing the cycle by connecting Hmp to R_{2m}. This simplification does not appear at first glance to be very different from that of the [$Hmp \cdot NO$, R_{6d}] edge. However, when we do this, we lose unidirectionality of the cycle in g_1 from Fig. 3.12 because [Hmp, R_1, $Hmp \cdot O_2$] is the only positive path in g_1. This causes the number of subgraphs of the corresponding fragment to explode: we gain four subgraphs analogous to g_1 but with different combinations of arrows, and one additional edge subgraph for a total of 10 subgraphs. The corresponding fragment is no longer critical—it has a coefficient of zero—illustrating the importance of maintaining directionality of cycles. Consequently the mechanism no longer has any critical fragments and saddle-node bifurcations are therefore no longer possible.

Chapter 4

Conclusions

4.1 Summary

(Bio)chemical networks have been frequent targets for simplification [21, 22]. In this thesis, I discussed graph-based dynamics-preserving schemes that can be applied to mass-action (bio)chemical mechanisms represented as bipartite graphs. In this representation, the graph contains two sets of vertices—one for species and one for reactions—and directed arrows (edges) from one type of vertex to the other. The theorems in this thesis preserve elements of bipartite graphs known as critical fragments, which are necessary for positive-feedback Andronov-Hopf and saddle-node bifurcations [20]. Because these are necessary conditions, critical fragment theory enables preliminary analysis of system dynamics. That is, the existence of a critical fragment is necessary for behaviors such as positive-feedback oscillations and multistationarity, whereas the absence of a critical fragment eliminates the possibility of these behaviors occurring near a steady state. Thus, it is imperative that reductions of models preserve critical fragments in order to preserve their dynamics.

In Sect. 3.3.1, I presented a set of theorems that will preserve a fragment's structure—and therefore its coefficient—based on the structure of the bipartite graph. This means that these rules can be applied to a model without knowing ahead of time what critical fragments it may have. These theorems were applied to a model for the control of the synthesis of the NO-detoxiying enzyme, Hmp, that displays bistability in Sect. 3.4.3. In Sect. 3.4.4, I discussed that knowledge of a system's critical fragments and its subgraphs may enable more aggressive reductions.

4.2 Discussion

4.2.1 Disadvantages

The graph-theoretical methods found in this thesis provide necessary, but not sufficient, conditions for positive-feedback oscillations and bistability [14, 19]. Hence, the preservation of these necessary conditions does not guarantee that the reduced model will preserve the original system's dynamics. However, even when dynamics is preserved, it is also possible for a (bio)chemical model to become disconnected from biology when applying the theorems and corollaries found in this thesis. This is due to the reduction rules being strictly graph-based without regard for a species' or reaction's role in the mechanism. Thus, it is highly recommended to identify species and reactions that are necessary to the mechanism's structure and function prior to any reductions, if possible.

4.2.2 Advantages

Perhaps the greatest advantage to the graph-based reduction rules presented in this thesis is their parameter independence. They require neither concentrations nor rate constants to be known, and all that is required is the set of elementary reactions that constitute the system. Thus, numerical considerations such as stiffness are irrelevant when applying these rules. Numerical methods, on the other hand, require known parameters and can be computationally taxing [21].

Another advantage to the graph-based reduction rules in this thesis is their capacity to simplify models without even knowing the model's critical fragments. This is because these rules preserve a fragment's coefficient post-reduction, i.e. a critical fragment would retain its criticality because its coefficient is invariant. Finding critical fragments is not exempt from being computationally taxing (particularly in larger networks); however, if critical fragments can be identified in the reduced model, it is simple to backtrack to the original set of reactions in the original model that would correspond to the vertices in the reduced critical fragment. Thus, these methods can enable us to find critical fragments that were

initially too large to enumerate by analyzing the reduced model's critical fragments.

4.3 Aggressive Reductions

The conservative reductions described in the theorems of Sect. 3.3.1 consist of a set of graph operations that preserve the coefficients by removing non-entry vertices. We also saw in Sect. 4.3 that, if we want to preserve specific, known critical fragments, and thus the associated dynamics, it may be possible to remove species and reactions intertwined with other parts of the network. These types of reductions are called *aggressive reductions*. Knowledge of critical fragments enables the potential deletion of entry vertices, such as free-floating edges in a critical fragment or intermediate edges along a cycle. However, we also saw in Sect. 3.4.4 that not all intermediate edges can be pruned. Here, we saw that removal of the $[R_1, Hmp \cdot O_2]$ product edge causes the cycle in g_1 from Fig. 3.12 to lose its unidirectionality because $[Hmp, R_1, Hmp \cdot O_2]$ is the only positive path in the cycle. This change consequently causes the critical fragment to lose its criticality, making saddle-node bifurcations, and thus bistability, no longer possible.

4.4 More General Transformations

As mentioned previously, this thesis discusses graph operations that preserve the coefficient of a fragment post-reduction. By extension, this means that if these graph-based reduction methods were applied within a critical fragment, then criticality of that fragment is preserved since the coefficient of the fragment under these rules is invariant. However, by definition, a critical fragment is one whose coefficient is negative. Theoretically, one can make transformations to a critical fragment that force its coefficient to become more negative or even enable a non-critical fragment to become critical by lowering its coefficient enough to drop below zero. The latter is potentially interesting because the genesis of critical fragments (in a system without any) could lead to *dynamics-inducing* schemes, in contrast to the reduction rules presented in this thesis that preserve the necessary conditions

for dynamics *already present* in the system. These types of transformations would typically be applied to linked cycles. As discussed in Sect. 3.2.2, reductions near an entry vertex may lead to the creation or destruction of other cycles, and these outcomes depend on the structure of the linked cycle and its surrounding vertices. To identify what effect changes near an entry vertex may have on the linked (or even isolated) cycle on any fragment it participates in, we would need to look at the fragment's subgraphs.

A fragment's subgraphs constitute the fragment. Once the subgraphs are enumerated, we can examine the ones with positive coefficients and deduce changes that can be made to lower some of their values and thus the value of the fragment. In this manner, the possibility of changes are endless. One possibility is by simply deleting reactant edges with large stoichiometric coefficients, or perhaps by shifting these coefficients to product edges since these edges do not appear in the edge subgraph. In the case of edge deletion, we would have to consider the reactant edge's participation in cycles (if any). Since reactant edges in the edge subgraph are squared in a coefficient's calculation, then it is possible that the magnitude of this value is greater than the magnitude of the value of the cycle's contribution to K_C that contains this edge, meaning that deletion of this edge would lower the fragment's coefficient and potentially turn a non-critical fragment critical. In the other case, if we shift a stoichiometric coefficient from a reactant edge to a product edge, then that edge's weight is no longer part of the edge subgraph, potentially leading to a decrease in a fragment's coefficient. Another possibility is by altering the parity of a cycle. If a cycle that appears in more than one subgraph has a net-positive contribution to K_C, removing a unitary negative path would change the value these subgraphs contribute to K_C to be net-negative, possibly granting criticality to a non-critical fragment. Put succinctly, further transformations can be constructed to force a fragment's coefficient to drop.

4.5 Future Directions

While critical fragment theory is effective in isolating key contributors to dynamics of (bio)chemical systems, it is not without its limitations. However, even without knowing the dynamics of a system, critical fragment theory can determine the *inability* of a system to express certain behaviors such as positive-feedback oscillations and multistationarity if the system contains no critical fragment of the required order [20]. Particularly with multistationarity, a critical fragment of order equal to the number of independent species is necessary. Since this is necessary and not sufficient, the system may not have the capacity for multistationarity even if it contains a critical fragment of the specified order. This uncertainty, however, may be reconciled by the union of critical fragment theory and other known theories.

Chemical Reaction Network Theory (CRNT) has similar applications in these systems and also utilizes graph-based analyses. One key difference in CRNT is that the species are noted as "complexes," which are defined as the combinations of chemical species at either end of a reaction arrow [13]. The graphs used in CRNT are not unlike our bipartite graphs; in particular, Species-Reaction (SR) graphs are indeed bipartite graphs with similar vertices: one for species and one for reactions. The edges, however, do not have direction and are labeled with the complex to which the species belongs in the reaction, in addition to the stoichiometric coefficient. In order to understand how CRNT and critical fragment theory can be complementary, we need to delve a bit further into the former's terminology.

In an SR graph, a "c-pair" (complex pair) is defined as a pair of edges that meet at a reaction vertex and have the same complex label (the complexes along the edges) [13]. A c-pair would correspond either to our notion of a negative path or to two products of a reaction. If a cycle has an even or odd number of c-pairs, it is an "e-cycle" or "o-cycle," respectively. An "s-cycle" is a cycle for which the result of alternately multiplying and dividing the coefficients along the cycle is 1. Lastly, two cycles have a "species-to-reaction intersection" (S-to-R intersection) if the common edges of the two cycles constitute a path

that begins at a species vertex and ends at a reaction vertex [13]. The simplest example of an S-to-R intersection in our terminology would be a single reactant edge shared by two cycles. The following is the main result in [13]:

Theorem 4 (Craciun and Feinberg [13])**.** *Consider a reaction network such that in its SR graph*

i each cycle is an o-cycle or an s-cycle,

ii no two e-cycles have an S-to-R intersection.

Then, taken with mass action kinetics, the reaction network does not have the capacity for multiple positive equilibria.

Violating at least one of these conditions is therefore a necessary condition for multistationarity. While the first condition does not have a direct correspondence with critical fragment theory, the second condition can indeed be translated using our nomenclature.[9] For example, if a system represented in our notion of a bipartite graph contains no linked cycles, or if a linked cycle's shared vertices do not form a reactant edge, then the second condition is met. Of course, this is not an easy condition to check for as it is likely a cycle would be linked in a bipartite graph, presumably with multiple shared vertices. However, due to the similarities with the construction of an SR graph and our bipartite graph, it would be very interesting to see how the two methods differ or are alike in the context of model reduction. Given that both seek to preserve key elements that determine the dynamic potential of a system, perhaps a future direction for graph-based reductions may lie in the junction of both CRNT and critical fragment theory.

[9]The first condition may be able to be determined if we label the edges in our bipartite graph with the complexes as well. But this would be for a future project.

Bibliography

[1] Martin Feinberg and Friedrich JM Horn. Dynamics of open chemical systems and the algebraic structure of the underlying reaction network. *Chem. Eng. Sci.*, 29(3):775–787, 1974.

[2] John J Tyson. Modeling the cell division cycle: cdc2 and cyclin interactions. *Proc. Natl. Acad. Sci. U.S.A.*, 88(16):7328–7332, 1991.

[3] Alan Mathison Turing. The chemical basis of morphogenesis. *Bull. Math. Biol.*, 52(1):153–197, 1990.

[4] AJ Koch and Hans Meinhardt. Biological pattern formation: from basic mechanisms to complex structures. *Rev. Mod. Phys.*, 66(4):1481, 1994.

[5] René Thomas and Morten Kaufman. Multistationarity, the basis of cell differentiation and memory. i. structural conditions of multistationarity and other nontrivial behavior. *Chaos*, 11(1):170–179, 2001.

[6] Martin Feinberg. The existence and uniqueness of steady states for a class of chemical reaction networks. *Arch. Ration. Mech. Anal.*, 132(4):311–370, 1995.

[7] Martin Feinberg. Multiple steady states for chemical reaction networks of deficiency one. *Arch. Ration. Mech. Anal.*, 132(4):371–406, 1995.

[8] GL Ermakov and BN Goldstein. Simplest kinetic schemes for biochemical oscillators. *Biochemistry (Moscow)*, 67(4):473–484, 2002.

[9] Maya Mincheva. Oscillations in biochemical reaction networks arising from pairs of subnetworks. *Bull. Math. Biol.*, 73(10):2277–2304, 2011.

[10] AN Ivanova. Conditions for uniqueness of the stationary states of kinetic systems, connected with the structures of their reaction-mechanisms. 1. *Kinet. Catal.*, 20(4):833–837, 1979.

[11] GL Ermakov. A theoretical graph method for search and analysis of critical phenomena in biochemical systems. i. graphical rules for detecting oscillators. *Biochemistry (Moscow)*, 68(10):1109–1120, 2003.

[12] Boris N Goldstein, Gennady Ermakov, Josep J Centelles, Hans V Westerhoff, and Marta Cascante. What makes biochemical networks tick? a graphical tool for the identification of oscillophores. *Eur. J. Biochem.*, 271(19):3877–3887, 2004.

[13] G Craciun and M Feinberg. Multiple equilibria in complex chemical reaction networks: II. The species-reaction graph. *SIAM J. Appl. Math.*, 66(4):1321–1338, 2006.

[14] Maya Mincheva and Marc R Roussel. Graph-theoretic methods for the analysis of chemical and biochemical networks. I. Multistability and oscillations in ordinary differential equation models. *J. Math. Biol.*, 55(1):61–86, 2007.

[15] Ovidiu Radulescu, Alexander N Gorban, Andrei Zinovyev, and Vincent Noel. Reduction of dynamical biochemical reactions networks in computational biology. *Front. Genet.*, 3:131, 2012.

[16] Stephen J Klippenstein, Vijay S Pande, and Donald G Truhlar. Chemical kinetics and mechanisms of complex systems: a perspective on recent theoretical advances. *J. Am. Chem. Soc.*, 136(2):528–546, 2014.

[17] Maya Mincheva and Marc R Roussel. A graph-theoretic method for detecting potential Turing bifurcations. *J. Chem. Phys.*, 125(20):204102, 2006.

[18] Marc R Roussel. *Nonlinear Dynamics: A hands-on introductory survey*. Morgan & Claypool Publishers, 2019.

[19] Carsten Conradi and Maya Mincheva. Graph-theoretic analysis of multistationarity using degree theory. *Math. Comput. Simul.*, 133:76–90, 2017.

[20] Georg R Walther, Matthew Hartley, and Maya Mincheva. Gratelpy: graph-theoretic linear stability analysis. *BMC Syst. Biol.*, 8(1):1–18, 2014.

[21] Thomas J Snowden, Piet H van der Graaf, and Marcus J Tindall. Methods of model reduction for large-scale biological systems: A survey of current methods and trends. *Bull. Math. Biol.*, 79(7):1449–1486, 2017.

[22] Alexander N Gorban. Model reduction in chemical dynamics: slow invariant manifolds, singular perturbations, thermodynamic estimates, and analysis of reaction graph. *Curr. Opin. Chem. Eng.*, 21:48–59, 2018.

[23] David J Klinke and Stacey D Finley. Timescale analysis of rule-based biochemical reaction networks. *Biotechnol. Prog.*, 28(1):33–44, 2012.

[24] Linda Petzold and Wenjie Zhu. Model reduction for chemical kinetics: An optimization approach. *AIChE J.*, 45(4):869–886, 1999.

[25] Olaf Wolkenhauer, Peter Wellstead, Kwang-Hyun Cho, and Brian Ingalls. Sensitivity analysis: from model parameters to system behaviour. *Essays Biochem.*, 45:177–194, 2008.

[26] Mutaz Khazaaleh, Sandhya Samarasinghe, and Don Kulasiri. A new hierarchical approach to multi-level model abstraction for simplifying ode models of biological networks and a case study: The g1/s checkpoint/dna damage signalling pathways of mammalian cell cycle. *Biosystems*, 203:104374, 2021.

[27] Th. de Donder. Sur la vitesse réactionelle. *Bull. Classe Sci. Acad. Roy. Belg.*, 23:936–941, 1937.

[28] I. Prigogine. Remarque sur le principe de réciprocité d'Onsager et le couplage des réactions chimiques. *Bull. Classe Sci. Acad. Roy. Belg.*, 32:30–35, 1946.

[29] Chitra Dutta, Anjan Dasgupta, and Jyotirmoy Das. On thermodynamic and kinetic equivalence of chemical systems. *Proc. Indian Acad. Sci. (Chem. Sci.)*, 93:817–830, 1984.

[30] Marc R Roussel. A delayed mass-action model for the transcriptional control of hmp, an no detoxifying enzyme, by the iron-sulfur protein FNR. pages 215–230, 2019.

[31] Jason C. Crack, Melanie R. Stapleton, Jeffrey Green, Andrew J. Thomson, and Nick E. Le Brun. Mechanism of [4Fe-4S](Cys)$_4$ cluster nitrosylation is conserved among NO-responsive regulators. *J. Biol. Chem.*, 288:11492–11502, 2013.

[32] Elizabeth A. M. Trofimenkoff and Marc R. Roussel. Small binding-site clearance delays are not negligible in gene expression modeling. *Math. Biosci.*, 325:108376, 2020.

[33] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási. The large-scale organization of metabolic networks. *Nature*, 407:651–654, 2000.

[34] David Kennell and Venugopal Talkad. Messenger RNA potential and the delay before exponential decay of messages. *J. Mol. Biol.*, 104:285–298, 1976.

[35] C. T. Walsh. Suicide substrates, mechanism-based enzyme inactivators: Recent developments. *Ann. Rev. Biochem.*, 53:493–535, 1984.

[36] Bard Ermentrout. *Simulating, Analyzing, and Animating Dynamical Systems*. SIAM, Philadelphia, 2002.

[37] Michail G. Slin'ko and Marina M. Slin'ko. Self-oscillations of heterogeneous catalytic reaction rates. *Catal. Rev. - Sci. Eng.*, 17:119–153, 1978.