AN ALL-ENCOMPASSING STABILITY RESULT FOR PROCESSIVE MULTISITE PHOSPHORYLATION SYSTEMS

MITCHELL EITHUN

Abstract. Phosphorylation systems are ubiquitous chemical mechanisms in biology. Multisite phosphorylation systems can be distributive or processive. Distributive systems have been shown to exhibit bistability, while processive systems exhibit global stability. However the processive result was proven for a specific mechanism of processive phosphorylation (namely, all catalytic reactions are reversible.) Accordingly, we generalize this result to allow for processive phosphorylation networks that are reversible or reversible or involve product inhibition. Specifically we create an all-encompassing processive system that encapsulates each of these schemes. By appealing to monotone systems theory we prove that the dynamical system arising from mass-action kinetics has a unique steady state and that it is a global attractor. We also establish the same result for a more general system using graph reductions and recent graph-theoretic stability criteria.

Keywords — Multisite phosphorylation, chemical reaction networks, mass-action kinetics, monotone systems theory, global stability, graph reductions

Contents

1. Introduction . . . 2
   1.1. Phosphorylation Mechanisms . . . 2
   1.2. More Details on Global Stability Results . . . 3
   1.3. Outline . . . 3
2. Background . . . 4
   2.1. Chemical Reaction Networks . . . 4
   2.2. Mass-Action Kinetics . . . 5
3. A Fully-Reversible Model . . . 6
   3.1. The ODE System . . . 7
4. Proof of Global Stability Using Monotone Systems Theory . . . 8
   4.1. Setup . . . 8
   4.2. Proof of Stability of the Fully-Reversible Model . . . 9
5. An All-Encompassing Model . . . 12
6. Proof of Global Stability Using Graph Reduction . . . 13
   6.1. Assumptions . . . 13
   6.2. Graph Constructions . . . 13
   6.3. Removing Intermediates . . . 15
   6.4. Stability Results . . . 17
   6.5. Proof of Global Stability of All-Encompassing Model . . . 18
7. Remarks . . . 19
Acknowledgments . . . 20
Appendices . . . 20

Date: July 21, 2016.

A. Bounded-Persistence 20
A.1. Definitions 20
A.2. Proof for the Reduced All-Encompassing Model 21
References 22

1. Introduction

1.1. **Phosphorylation Mechanisms.** A biological process of great importance, *phosphorylation* is the enzyme-mediated addition of a phosphate group to a protein substrate, which often modifies the function of the substrate. This basic mechanism is: $S_0+E \leftrightarrow S_0E \to S_1+E$, where S_i is the substrate with i phosphate groups attached and E is the enzyme.

Additionally, many substrates have more than one *site* at which phosphate groups can be attached. Such multisite phosphorylation may be *distributive* or *processive*, or somewhere in between [10, 14]. In distributive phosphorylation, each binding of substrate and enzyme results in at most one addition of a phosphate group. In contrast, in processive phosphorylation, when an enzyme catalyzes the addition of a phosphate group, phosphate groups are added to all sites before the enzyme and substrate dissociate.

Most studies on the mathematics of multisite phosphorylation have focused on multisite phosphorylation under a sequential and fully *distributive* mechanism. These systems admit bistability [12] and oscillations [9].

As for *processive* phosphorylation, in [6], Conradi and Shiu considered the following processive n-site phosphorylation/dephosphorylation network (also called the "futile cycle"):

$$\begin{aligned} &S_0 + K \underset{k_2}{\overset{k_1}{\rightleftharpoons}} S_0K \underset{k_4}{\overset{k_3}{\rightleftharpoons}} S_1K \underset{k_6}{\overset{k_5}{\rightleftharpoons}} \dots \underset{k_{2n}}{\overset{k_{2n-1}}{\rightleftharpoons}} S_{n-1}K \xrightarrow{k_{2n+1}} S_n + K \\ &S_n + F \underset{\ell_{2n}}{\overset{\ell_{2n+1}}{\rightleftharpoons}} S_nF \underset{\ell_{2n-2}}{\overset{\ell_{2n-1}}{\rightleftharpoons}} \dots \underset{\ell_4}{\overset{\ell_5}{\rightleftharpoons}} S_2F \underset{\ell_2}{\overset{\ell_3}{\rightleftharpoons}} S_1F \xrightarrow{\ell_1} S_0 + F \end{aligned} \tag{1}$$

and proved that the resulting systems, in contrast with distributive systems, do *not* admit bistability or oscillations, and, moreover, exhibit rigid dynamics: each invariant set contains a unique steady state, which is a global attractor [6]. This result was proven by generalizing a result of [2]. Using other means, Rao [15] and Marcondes de Freitas, et al. [8] have established the same result.

However, there are other possible mechanisms for processive phosphorylation, the following being the most common [17]:

$$\begin{aligned} &S_0 + K \underset{k_2}{\overset{k_1}{\rightleftharpoons}} S_0K \xrightarrow{k_3} S_1K \xrightarrow{k_5} \dots \xrightarrow[k_{2n}]{k_{2n-1}} S_{n-1}K \xrightarrow{k_{2n+1}} S_n + K \\ &S_n + F \underset{\ell_{2n}}{\overset{\ell_{2n+1}}{\rightleftharpoons}} S_nF \xrightarrow{\boldsymbol{\ell_{2n-1}}} \dots \xrightarrow{\ell_5} S_2F \xrightarrow{\ell_3} S_1F \xrightarrow{\ell_1} S_0 + F \end{aligned} \tag{2}$$

Here, only the first reaction is reversible.

As another example, we incorporate *product inhibition*, in which an intermediate is added before the substrate is form in each component. A processive realization of this scheme is

$$\begin{gathered} S_0 + K \rightleftharpoons S_0K \longrightarrow S_1K \longrightarrow \ldots \longrightarrow S_{n-1}K \longrightarrow S_nK \rightleftharpoons S_n + K \\ S_n + F \rightleftharpoons S_nF \longrightarrow \ldots \longrightarrow S_2F \longrightarrow S_1F \longrightarrow S_0F \underset{\ell^*}{\rightleftharpoons} S_0 + F \end{gathered} \tag{3}$$

There are distributive systems in the literature with such product inhibition [13, Scheme 2]. The final reaction in each of the two components is inhibited via a back-reaction.

Can the global stability result for (1) be generalized to incorporate the other mechanisms (2) and (3)? We accomplish this in this paper (Theorem 3.2) by introducing a *fully-reversible* system which encapsulates (1), (2), (3) We prove that the dynamical system associated with this system is globally stable. This result generalize results of Angeli and Sontag [2], Conradi and Shiu [6], Rao [15] and Marcondes de Freitas, et al. [8]

1.2. **More Details on Global Stability Results.** Chemical reaction networks, such as phosphorylation, are represented as directed graphs that connect chemical species. Using *mass-action kinetics*, the dynamics of the system can be realized as a system of ODEs. At *steady state* concentration, each ODE in the system vanishes and so the concentration of each chemical species is constant. A primary focus of chemical reaction network research is to understand the *stability* of steady states, that is whether or not concentrations tend to approach a steady state. If all concentrations in some neighborhood of the steady state approach the steady state as $t \to \infty$, a steady state is *locally asymptotically stable*.

Many authors, including [5] and [15], have established that the 1-site phosphorylation has a unique steady state, and it is *globally asymptotically stable*, meaning that any trajectory starting in the positive orthant converges to the steady state as $t \to \infty$. As mentioned above, this result was generalized by Angeli and Sontag who showed that that the *processive* n-site system (1) admits a unique steady state and it is globally stable [6].

We further generalize this result by establishing that our fully-reversible network is globally stable. The first proof extends the result in [6], which applies results from monotone systems theory from [2]. This involves assuming a change of coordinates, manipulating matrices and computing the Jacobian matrix of the system and applying graph-theoretic criteria.

We also construct a second network, the *all-encompassing* model, which further generalizes the fully-reversible model, allowing for m reaction components, rather than 2. Furthermore, each component is allowed to have a different number of binding sites. We prove that this model is globally stable using recent graph-theoretic criteria developed by Marcondes de Freitas, et al. in [1] and [8]. Using the scheme in [1], the chemical reaction system is used to construct two labeled graphs, an R-graph and SR-graph. If these graphs satisfy certain conditions, the chemical system is globally stable. In [8] it was shown that we can perform a reduction on the original system by removing intermediate species before using the stability criterion. This makes for a faster proof of global stability. Also, this new result generalizes a recent global stability result due to Rao [15].

1.3. **Outline.** In section 2 we define a chemical reaction network and explain how mass-action kinetics produces an associated dynamical systems. In section 3, we develop our fully-reversible n-site, 2-component network, which captures systems that are reversible, irreversible and have product inhibition. The global stability of the system is established

in section 4. Section 5 explains the all-encompassing m-component model and in section 6 we prove it is globally stable using graph-theoretic reductions. To conclude we comment on how and our models compare with other phosphorylation systems in section 7. Appendix A provides some a brief explanation of how to establish the technical detail of bounded persistence.

2. Background

This section provides an overview of how mass-action kinetics produces a dynamical system for a given chemical reaction network. Our setup is based on [6] and [8].

2.1. Chemical Reaction Networks.

As an example, consider the chemical reaction

$$A + B \xrightarrow{\kappa} 3A + C \tag{4}$$

The graph of (4) is a *chemical reaction network*. The vertices $A + B$ and $3A + C$ are *complexes*, which are linear combinations of individual *species*. The complex on the left side of a reaction is called a *reactant* and the complex on the right side of a reaction is a *product*.

An *irreversible* reaction is denoted by a unidirectional arrow ($\to$). A reaction with a double arrow, such as $X \leftrightharpoons Y$ denotes a *forward reaction* $X \to Y$ and a *backward reaction* $Y \to X$. Together these reactions are known as a *reversible* reaction. The parameter κ is known as a *rate constant*. In the case of (4), one unit of A and one unit of B react at rate proportional to their concentrations, with constant of proportionality κ, to form three units of A and one unit of C.

More formally, we express as G chemical reaction network with n species as the triple $G = (\mathcal{S}, \mathcal{C}, \mathcal{R})$, which consists of

(1) a finite nonempty set of species $\mathcal{S} = S_1, \ldots, S_n$,
(2) a set of complex vectors $\mathcal{C}$ of the form $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n_{\geq 0}$, representing the weights on a linear combination species, and
(3) a set of reversible ($X \leftrightharpoons Y$) and irreversible ($X \to Y$) reactions $\mathcal{R}$

Consider the species $y = (\alpha_1, \ldots, \alpha_n)$ and $y' = (\alpha'_1, \ldots, \alpha'_n)$. For a reaction $y \to y'$ or $y \leftrightharpoons y'$, we call $y - y'$ the *reaction vector*, which describes the net change in species.

In biology, *phosphorylation* is a chemical mechanism that adds a phosphate group. The 1-site phosphorylation is shown below.

Example 2.1. The following network (called the "futile cycle") describes 1-site phosphorylation:

$$S_0 + K \underset{k_2}{\overset{k_1}{\rightleftharpoons}} S_0K \xrightarrow{k_3} S_1 + K \tag{5}$$

$$S_1 + F \underset{\ell_2}{\overset{\ell_3}{\rightleftharpoons}} S_1F \xrightarrow{\ell_1} S_0 + F$$

The key players in this network are a kinase (K), a phosphatase (F), and a substrate (S_0). The substrate S_1 is obtained from the unphosphorylated protein S_0 by attaching a phosphate group to it via an enzymatic reaction catalyzed by K. Conversely, a reaction catalyzed by F removes the phosphate group from S_1 to obtain S_0. The intermediate complexes S_0K and S_1F are the bound enzyme-substrate complexes.

2.2. **Mass-Action Kinetics.** Consider the example (4). Let x_A, x_B and x_C be the concentrations of the species as functions of time. Assuming the reaction follows *mass-action kinetics*, the species A and B react proportionally to the product of their concentrations with constant of proportionally κ. Noting that the reaction yields a net change of two units in the amount of A, we obtain the first differential equation in the following system:

$$\begin{aligned} \frac{d}{dt}x_A &= 2\kappa x_A x_B \\ \frac{d}{dt}x_B &= -\kappa x_A x_B \\ \frac{d}{dt}x_C &= \kappa x_A x_B \ . \end{aligned}$$

The other equations follow similarly. The mass-action differential equations that a network defines are comprised of a sum of the monomial contribution from the reactant of each chemical reaction in the network; these differential equations will be defined by equations (6–7).

Letting m denotes the number of reactions, where we count each pair of reversible reactions only once, the *stoichiometric matrix* Γ is the $s \times m$ matrix whose k-th column is the reaction vector of the k-th reaction (in the forward direction if the reaction is reversible), i.e., it is the reaction vector $y_j - y_i$ if k indexes the (forward) reaction $y_i \to y_j$. The choice of kinetics is encoded by a locally Lipschitz function $R : \mathbb{R}^s_{\geq 0} \to \mathbb{R}^m$ that encodes the reaction rates of the m reactions as functions of the s species concentrations (a pair of reversible reactions is counted only once – in this case, R_k is the forward rate minus the backward rate). The *reaction kinetics system* defined by a reaction network G and reaction rate function R is given by the following system of ODEs:

$$\frac{dx}{dt} = \Gamma\, R(x) \ . \tag{6}$$

For *mass-action kinetics*, which is the setting of this paper, the coordinates of R are:

$$R_k(x) = \begin{cases} \kappa_{ij} x^{y_i} & \text{if } k \text{ indexes an irreversible reaction } y_i \to y_j \\ \kappa_{ij} x^{y_i} - \kappa_{ji} x^{y_j} & \text{if } k \text{ indexes a reversible reaction } y_i \leftrightarrow y_j \end{cases} \tag{7}$$

A *chemical reaction system* refers to the dynamical system (6) arising from a specific chemical reaction network G and a choice of rate parameters $(\kappa^*_{ij}) \in \mathbb{R}^r_{>0}$ (recall that r denotes the number of reactions) where the reaction rate function R is that of mass-action kinetics (7).

The *stoichiometric subspace* is the vector subspace of $\mathbb{R}^s$ spanned by the reaction vectors $y_j - y_i$ (where (i, j) is an edge of G), and we will denote this space by $\mathcal{S}$:

$$\mathcal{S} := \mathbb{R}\{y_j - y_i \mid (i, j) \in E\} \ . \tag{8}$$

Note that in the setting of (6), one has $\mathcal{S} = \mathrm{im}(\Gamma)$. In the earlier example reaction shown in (4), we have $y_2 - y_1 = (2, -1, 1)$, which means that with each occurrence of the reaction, two units of A and one of C are produced, while one unit of B is consumed. This vector $(2, -1, 1)$ spans the stoichiometric subspace $\mathcal{S}$ for the network (4). Note that the vector $\frac{dx}{dt}$ in (6) lies in $\mathcal{S}$ for all time t. In fact, a trajectory $x(t)$ beginning at a positive vector $x(0) = x^0 \in \mathbb{R}^s_{>0}$ remains in the *stoichiometric compatibility class* (also called an "invariant

polyhedron"), which we denote by

$$\mathcal{P} \;:=\; (x^0 + \mathcal{S}) \cap \mathbb{R}^s_{\geq 0}\;, \tag{9}$$

for all positive time. In other words, this set is forward-invariant with respect to the dynamics (6). A *steady state* of a reaction kinetics system (6) is a nonnegative concentration vector $x^* \in \mathbb{R}^s_{\geq 0}$ at which the ODEs (6) vanish: $\Gamma R(x^*) = 0$. We distinguish between *positive steady states* $x^* \in \mathbb{R}^s_{>0}$ and *boundary steady states* $x^* \in \left(\mathbb{R}^s_{\geq 0} \setminus \mathbb{R}^s_{>0}\right)$. A system is *multistationary* (or *admits multiple steady states*) if there exists a stoichiometric compatibility class $\mathcal{P}$ with two or more positive steady states. In the setting of mass-action kinetics, a network may admit multistationarity for all, some, or no choices of positive rate constants κ_{ij}.

For an example of how mass-action kinetics generates a dynamical system, see [6], which explains how to write-down the ODE system (6) for the 1-site phosphorylation system in Example 2.1.

3. A Fully-Reversible Model

In this section we introduce a generalized version of the n-site processive phosphorylation system, which we call the **fully-reversible** model. This system captures different variants on processive phosphorylation. In contrast to system (1), the system studied in [6], we allow the final concentrations of $S_n + K$ and $S_0 + F$ to be inhibited by rate constants k^* and ℓ^* and we allow rate constants of back-reactions to be zero. The reaction network for our model is shown below.

$$\begin{aligned} S_0 + K &\underset{k_2}{\overset{k_1}{\rightleftharpoons}} S_0K \underset{k_4}{\overset{k_3}{\rightleftharpoons}} S_1K \underset{k_6}{\overset{k_5}{\rightleftharpoons}} \dots \underset{k_{2n}}{\overset{k_{2n-1}}{\rightleftharpoons}} S_{n-1}K \underset{k^*}{\overset{k_{2n+1}}{\rightleftharpoons}} S_n + K \\ S_n + F &\underset{\ell_{2n}}{\overset{\ell_{2n+1}}{\rightleftharpoons}} S_nF \underset{\ell_{2n-2}}{\overset{\ell_{2n-1}}{\rightleftharpoons}} \dots \underset{\ell_4}{\overset{\ell_5}{\rightleftharpoons}} S_2F \underset{\ell_2}{\overset{\ell_3}{\rightleftharpoons}} S_1F \underset{\ell^*}{\overset{\ell_1}{\rightleftharpoons}} S_0 + F \end{aligned} \tag{10}$$

where

$$\begin{aligned} k_2, k_4, \dots, k_{2n}, \ell_2, \ell_4, \dots, \ell_{2n}, k^*, \ell^* &\geq 0, \\ k_1, k_3, \dots, k_{2n+1}, \ell_1, \ell_3, \dots, \ell_{2n+1} &> 0. \end{aligned}$$

This model is general enough that it reduces to several well-defined systems under certain conditions, leading to the following proposition, with systems numbered as they appear in section 1.

Proposition 3.1. *The fully-reversible model* (10) *encompasses the following phosphorylation systems:*

(1) Reversible systems, which allow for binding in both directions.
(2) Irreversible systems, which have no back-reactions after initial binding.
(3) Systems with product inhibition.

The conditions that we impose on (10) to find these variations are listed in Table 1.

System variant	Conditions
$(n-1)$-site system with product inhibition	$n \geq 2$ and $k^*, \ell^* > 0$
Reversible n-site system	$\ell^*, k^* = 0$ and $k_2, k_4, \ldots, k_{2n}, \ell_2, \ell_4, \ldots, \ell_{2n} > 0$
Irreversible n-site system	$\ell^*, k^* = 0$

TABLE 1. The restrictions on rate constants that reduce the all-encompassing model to different variants of processive phosphorylation systems.

Remark 1. Notice that the reversible system described in Table 1 is equivalent to the n-site system (1), studied in [6]. Clearly the model (10) also captures other theoretical families of distributive sessions, including, for example, irreversible n-site systems with product inhibition. In section 7 we remark on other types of phosphorylation systems.

Our first main theorem is that this system is globally stable. It follows from Proposition 3.1 that the systems described in Table 1 are globally stable.

Theorem 3.2. *The dynamical system* (6) *of the all-encompassing model* (10) *arising from mass-action kinetics has a unique positive steady state and it is a global attractor.*

We prove this result in Section 4.

3.1. **The ODE System.** The variables in the fully reversible mode (10) are ordered according to Table 2.

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	$\cdots$	x_{2n+3}	x_{2n+4}
K	F	S_0	S_n	S_0K	S_1F	S_1K	S_2F	$\cdots$	$S_{n-1}K$	S_nF

TABLE 2. Assignment of variables and species of the processive n-site network (10)

Using mass-action kinetics we can concisely express the system as the matrix equation (6). The stoichiometric matrix, whose columns span the stoichiometric subspace, is given by the $(2n+2) \times (2n+4)$ matrix

$$\begin{aligned} \Gamma = \big[e_5 - (e_1 + e_3) \quad | \quad \ldots, e_{2i+5} - e_{2i+3}, \ldots \quad | \quad e_4 + e_1 - e_{2n+3}, \\ e_2 + e_3 - e_6 \quad | \quad \ldots, e_{2i+4} - e_{2i+6}, \ldots \quad | \quad e_{2n+4} - (e_2 + e_4) \big] \end{aligned} \tag{11}$$

where $i = 1, \ldots, n-1$. The stoichiometric matrix (and, by extension, the stoichiometric subspace) is the same as in [6]. The following Lemma is proved in [6]:

Lemma 3.3. *The stoichiometric matrix* Γ *for* (10) *has rank* $2n-1$.

Remark 2 (Conservation relations). By Lemma 3.3, $\ker(\Gamma^t) = \mathcal{S}^\perp$ is three-dimensional. A particular basis is formed by the rows of the following matrix:

$$\mathcal{A} = \left[\begin{array}{cc|cc|cccccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & \cdots & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & \cdots & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 & 1 \end{array}\right]. \tag{12}$$

This basis has the following interpretation: the total amounts of free and bound enzyme or substrate remain constant as the dynamical system (6) progresses. In other words, the rows

of $\mathcal{A}$ correspond to the following conserved (positive) quantities (recall the species ordering from Table 2):

$$\begin{aligned} K_{\text{tot}} &= x_1 + (x_5 + x_7 + \cdots + x_{2n+3}), \\ F_{\text{tot}} &= x_2 + (x_6 + x_8 + \cdots + x_{2n+4}) \ , \\ S_{\text{tot}} &= x_3 + x_4 + \cdots + x_{2n+4} \ . \end{aligned}$$

The reaction rate function arising from mass-action kinetics for (10) is,

$$R(x) = \begin{bmatrix} k_1 x_1 x_3 - k_2 x_5 \\ \hline k_3 x_5 - k_4 x_7 \\ k_5 x_7 - k_6 x_9 \\ \vdots \\ k_{2n-1} x_{2n+1} - k_{2n} x_{2n+3} \\ \hline k_{2n+1} x_{2n+3} - k^* x_1 x_4 \\ \hline \ell_{2n+1} x_2 x_4 - \ell_{2n} x_{2n+4} \\ \hline \ell_{2n-1} x_{2n+4} - \ell_{2n-2} x_{2n+2} \\ \ell_{2n-3} x_{2n+2} - \ell_{2n-4} x_{2n} \\ \vdots \\ \ell_3 x_8 - \ell_2 x_6 \\ \hline \ell_1 x_6 - \ell^* x_2 x_3 \end{bmatrix} . \tag{13}$$

Using the stoichiometric matrix (11) and the reaction rate function (13) we can write the dynamical system arising from the fully-reversible system (10) as the equation (6). We appeal to this formulation to prove global stability in the next section.

4. Proof of Global Stability Using Monotone Systems Theory

In this section, we prove that each steady state of the processive network taken with mass-action kinetics is a global attractor of the corresponding compatibility class (Theorem 4.2). The argument uses the same scheme as in [6]. Much of section 4.1 is adapted from [6] and the proof in section 4.2 extends that in [6, §6].

4.1. **Setup.** We begin by recalling the setup in Angeli and Sontag [2, §3]. We consider any reaction kinetics system with s chemical species and m reactions (where each pair of reversible reactions is counted only once) given by $\dot{x} = \Gamma\, R(x)$, as in (6). Each such system together with a vector $\sigma \in \mathbb{R}^s_{\geq 0}$ (viewed as an initial condition of (6)) defines another ODE system:

$$\dot{c} = f_\sigma(c) := R(\sigma + \Gamma c), \tag{14}$$

with associated state space (which is sometimes called the space of "reaction coordinates")

$$X_\sigma = \left\{ c \in \mathbb{R}^m \mid \sigma + \Gamma c \in \mathbb{R}^s_{\geq 0} \right\}. \tag{15}$$

For simplicity, we introduce $z := \sigma + \Gamma c$. For $i = 1, \ldots, n$, define

$$Z_i := \begin{bmatrix} z_{2i+3} \\ z_{2i+4} \end{bmatrix} = \begin{bmatrix} \sigma_{2i+3} & + & c_i & - & c_{i+1} \\ \sigma_{2i+4} & + & c_{2i+2} & - & c_{2i+1} \end{bmatrix}.$$

Using the stoichiometric matrix (11), we have

$$z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix} = \begin{bmatrix} \sigma_1 + c_{n+1} - c_1 \\ \sigma_2 + c_{n+2} - c_{2n+2} \\ \sigma_3 + c_{n+2} - c_1 \\ \sigma_4 + c_{n+1} - c_{2n+2} \\ Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix}. \tag{16}$$

To state Lemma 4.1 below, we require the following definition.

Definition 4.1.

(1) The nonnegative orthant $\mathbb{R}^m_{\geq 0}$ defines a partial order on $\mathbb{R}^m$ given by $c_1 \succcurlyeq c_2$ if $c_1 - c_2 \in \mathbb{R}^m_{\geq 0}$. Also, we write $c_1 \succ c_2$ if $c_1 \succcurlyeq c_2$ with $c_1 \neq c_2$, and $c_1 \gg c_2$ if $c_1 - c_2 \in \mathbb{R}^m_{>0}$.

(2) A dynamical system with state space $X \subseteq \mathbb{R}^m$ and flow denoted by $\phi_t(c)$ (for initial condition c) is *monotone with respect to the nonnegative orthant* $\mathbb{R}^m_{\geq 0}$ if the partial order arising from $\mathbb{R}^m_{\geq 0}$ is preserved by the forward flow: for $c_1, c_2 \in X$, if $c_1 \geq c_2$ then $\phi_t(c_1) \geq \phi_t(c_2)$ for all $t \geq 0$. A dynamical system is *strongly monotone with respect to the nonnegative orthant* if it is monotone with respect to the nonnegative orthant and, additionally, for $c_1, c_2 \in X$, the relation $c_1 \succ c_2$ implies that $\phi_t(c_1) \gg \phi_t(c_2)$ for all $t > 0$.

The lemma below is due to Angeli and Sontag [2, Corollary 3.3]. We note that it is stated in the setting of monotonicity with respect to the nonnegative orthant (cone), but the result and the theory of monotone systems more generally extend to other cones and moreover to partial orders not necessarily arising from a cone [4].

The following lemma was originally proven by Angeli and Sontag in [2, Corollary 3.3] and appeals to monotone systems theory to establish the presence of a unique, globally stable steady state.

Lemma 4.1 (Angeli and Sontag)**.** *Let R, Γ, and σ be as in the setup above. Assume that:*

(1) the stoichiometric matrix Γ has rank $m-1$, with kernel spanned by some positive vector (i.e., in $\mathbb{R}^m_{>0}$),

(2) every trajectory of the reaction kinetics system (6) is bounded, and

(3) the system $\dot{c} = f_\sigma(c)$ defined in (14) is strongly monotone with respect to the nonnegative orthant.

Then there exists a unique $\eta = \eta_\sigma \in \mathbb{R}^s_{\geq 0}$ such that for any initial condition $\mu \in \mathbb{R}^s_{\geq 0}$ that is stoichiometrically compatible with σ (i.e., $\mu - \sigma \in \mathrm{Im}(\Gamma)$), the trajectory $x(t)$ of the reaction kinetics system (6) with initial condition $x(0) = \mu$ converges to η: $\lim_{t\to\infty} x(t) = \eta$.

In [6], Lemma 4.1 is used to prove that the n-site processive system is globally stable provided all rate constants are strictly positive. We expand the proof to establish the global stability of the fully-reversible model (10).

4.2. **Proof of Stability of the Fully-Reversible Model.** We restate Theorem 3.2 more precisely below.

Theorem 4.2. *Let n be a positive integer. For any chemical reaction system* (6) *arising from the processive n-site network* (10) *and any choice of rate constants,*

(1) each stoichiometric compatibility class $\mathcal{P}$ contains a unique steady state η,
(2) η is a positive steady state, and
(3) η is the global attractor of $\mathcal{P}$.

Proof. Let $\sigma \in \mathcal{P}$. The result will follow from Lemma 4.1 applied to this reaction system and the vector σ, once we verify its three hypotheses.

For hypothesis (1) we note that the rank of Γ is $(2n+2)-1$ by (3.3).

For hypothesis (2) of Lemma 4.1, every stoichiometric compatibility class is bounded due to the conservation laws. Thus, trajectories of (6) are bounded.

Finally, we must verify that the system (14) is strongly monotone. We begin by showing that it is monotone with respect to the nonnegative orthant. It suffices (by Proposition 1.1 and Remark 1.1 in [16, §3.1]) to show that the Jacobian matrix of $f_\sigma(c) := R(\sigma + \Gamma c)$ with respect to c has nonnegative off-diagonal entries for all $c \in X_\sigma$. Note that this reaction rate function R appeared earlier in (13).

For simplicity, we introduce $z := \sigma + \Gamma c$, so by the chain rule, the Jacobian matrix of $f_\sigma(c) := R(\sigma + \Gamma c)$ with respect to c is $\mathrm{Jac}_c f_\sigma(c) = \mathrm{Jac}_x R(z)\,\Gamma$, which is the following $(2n+2)\times(2n+2)$-matrix:

$$\begin{bmatrix}
(-k_1(z_3+z_1)-k_2)\,e_1 + k_2 e_2 + k_1 z_3 e_{n+1} + k_1 z_1 e_{2n+2} \\
\hline
k_3 e_1 - (k_3+k_4)e_2 + k_4 e_3 \\
k_5 e_2 - (k_5+k_6)e_3 + k_6 e_4 \\
\vdots \\
k_{2n-1} e_{n-1} - (k_{2n-1}+k_{2n})e_n + k_{2n} e_{n+1} \\
\hline
k_{2n+1} e_n - k_{2n+1} e_{n+1} - k^* z_1 (e_{n+1} - e_{2n+2}) - k^* z_4 (e_{n+1} - e_1) \\
\hline
\ell_{2n+1} z_2 e_{n+1} + (-\ell_{2n+1}(z_4+z_2) - \ell_{2n})\,e_{n+2} + \ell_{2n} e_{n+3} + \ell_{2n+1} z_4 e_{n+2} \\
\hline
\ell_{2n-1} e_{n+2} - (\ell_{2n-1}+\ell_{2n-2})e_{n+3} + \ell_{2n-2} e_{n+4} \\
\ell_{2n-3} e_{n+3} - (\ell_{2n-3}+\ell_{2n-4})e_{n+4} + \ell_{2n-4} e_{n+5} \\
\vdots \\
\ell_3 e_{2n} - (\ell_3+\ell_2)e_{2n+1} + \ell_2 e_{2n+2} \\
\hline
\ell_1 e_{2n+1} - \ell_1 e_{2n+2} - \ell^* z_2 (e_{n+2} - e_1) - \ell^* z_3 (e_{n+2} - e_{2n+2})
\end{bmatrix}. \tag{17}$$

By inspection of the Jacobian matrix (17), each nonzero off-diagonal entry either is some ℓ_i or k_j, which is nonnegative, or has the form $k_j z_i$ or $\ell_j z_i$ (for some i) and such a term is nonnegative for $c \in X_\sigma$ (recall that the system (14) evolves on the space X_σ defined in (15), so $z = \sigma + \Gamma c \in \mathbb{R}^{2n+2}_{\geq 0}$.

Now we show that the system (14) is strongly monotone by checking that the Jacobian matrix (17) is almost everywhere irreducible along trajectories of (14) (see Theorem 1.1 of [16, §4.1]), i.e., that the matrix is almost everywhere the adjacency matrix of a strongly connected directed graph. By inspection of (17), this directed graph always contains the paths $n+1 \to n \to \cdots \to 1$ and $2n+2 \to 2n+1 \to \cdots \to n+2$, and the only possible edges between these two components are $1 \to 2n+2$ and $n+2 \to n+1$, so we must show that the corresponding two entries in the matrix (17), namely $k_1 z_1 = k_1(\sigma_K - c_1 + c_{n+1})$ and $\ell_{2n+1} z_2 = \ell_{2n+1}(\sigma_F - c_{n+2} + c_{2n+2})$, are almost everywhere nonzero along trajectories. The pertinent edges are illustrated in Figure 1.

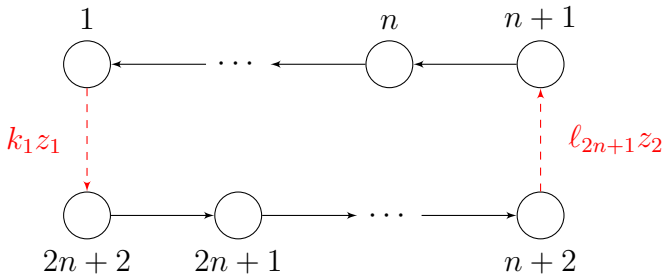


FIGURE 1. Some of the edges in the directed graph for which the adjacency matrix is the Jacobian matrix. To show that the graph is strongly connected we need show that $z_1(t)$ and $z_2(t)$ are almost everywhere nonzero.

By symmetry between K and F, we need only verify that $z_1(t)$ is almost everywhere nonzero along trajectories. Suppose, on the contrary, that $z_1(t) \equiv 0$ on some non-degenerate time interval. By the conservation law on the total amount of Kinase, the following sum must be positive:

$$\begin{aligned}\sigma_1 + (\sigma_5 + \cdots + \sigma_{2n+3}) &= (\sigma_1 - c_1 + c_{n+1}) + [(\sigma_5 + c_1 - c_2) + \cdots + (\sigma_{2n+3} + c_n - c_{n+1})] \\ &= z_1(t) + [z_5(t) + \cdots + z_{2n+3}(t)]. \end{aligned} \tag{18}$$

By assumption, $z_1(t) \equiv 0$, so if we show that $z_5(t), \ldots, z_{2n+3}(t) \equiv 0$, we have a contradiction.

From (16) we have $z_1(t) = \sigma_1 - c_1(t) + c_{n+1}(t)$. Taking the derivative implies

$$\begin{aligned} 0 &\equiv -\dot{c}_1(t) + \dot{c}_{n+1}(t) \\ &= -[k_1 z_1(t) z_3(t) - k_2 z_5(t)] + [k_{2n+1} z_{2n+3}(t) - k^* z_1(t) z_4(t)] \\ &= k_2 z_5(t) + k_{2n+1} z_{2n+3}(t). \end{aligned}$$

We know that $k_{2n+1} > 0$ and the $k_2, z_5(t), z_{2n+3}(t) \geq 0$ and so, for the above statement to hold, it must be true that $z_{2n+3}(t) \equiv 0$. If $n = 1$, we are done. Otherwise, we have

$$0 \equiv z_{2n+3}(t) = \sigma_{2n+3} + c_n - c_{n+1}. \tag{19}$$

Taking the derivative of (19) implies

$$\begin{aligned} 0 &\equiv \dot{c}_n - \dot{c}_{n+1} \\ &= [k_{2n-1} z_{2n+1}(t) - k_{2n} z_{2n+3}(t)] - [k_{2n+1} z_{2n+3}(t) - k^* z_2(t) z_4(t)] \\ &= -k_{2n-1} z_{2n+1}(t). \end{aligned}$$

We know that $k_{2n-1} > 0$ and so it follows that $z_{2n+1}(t) \equiv 0$.

Next we prove inductively that for $i = 0, \ldots, n-1$,

$$z_{2(n-i)+3}(t) \equiv 0. \tag{20}$$

We have established that (20) holds for $i = 0$ and $i = 1$ so we are done if $n = 2$. Otherwise this is the base case. For the induction hypothesis, suppose that (20) holds for $i = k-2$ and $i = k-1$, where $2 \leq k \leq n-1$. This means $z_{2(n-k)+5}(t) \equiv 0$ and $z_{2(n-k)+7}(t) \equiv 0$. By construction,

$$0 \equiv z_{2(n-(k-1))+3}(t) = z_{2(n-k)+5}(t) = \sigma_{2(n-k)+5} + c_{n-k+1} - c_{n-k+2}. \tag{21}$$

Taking the derivative of (21) implies

$$\begin{aligned}0 &\equiv \dot{c}_{n-k+1} - \dot{c}_{n-k+2}\\ &= k_{2(n-k)+1}z_{2(n-k)+3}(t) - k_{2(n-k)+2}z_{2(n-k)+5}(t) - [k_{2(n-k)+3}z_{2(n-k)+5}(t) - k_{2(n-k)+4}z_{2(n-k)+7}(t)]\\ &= k_{2(n-k)+1}z_{2(n-k)+3}(t).\end{aligned}$$

Because $k_{2(n-k)+1} > 0$, we can conclude that $z_{2(n-k)+3}(t) \equiv 0$.

We have shown that $z_5(t), \ldots, z_{2n+3}(t) \equiv 0$, which contradicts the the fact that the conservation law (18) must be positive. It follows that $z_1(t)$ is almost everywhere nonzero along trajectories, and the proof is complete. □

5. An All-Encompassing Model

Next we consider a more general class of phosphorylation systems in which there are more than two substrates and enzymes. We call this the all-encompassing model and it generalizes the fully-reversible model in two ways:

(1) There are now m components rather than 2, each with its own enzyme E_i and substrate P_i.
(2) Each of the m components has n_i binding sites rather than a fixed n.

The **all-encompassing** reaction network for our model is shown below.

$$\begin{array}{ccccccc}
P_1 + E_1 & \underset{k_{-11}}{\overset{k_{11}}{\rightleftarrows}} & C_{11} & \underset{k_{-12}}{\overset{k_{12}}{\rightleftarrows}} & C_{12} & \underset{k_{-13}}{\overset{k_{13}}{\rightleftarrows}} \quad \cdots \quad \underset{k_{-1n_1}}{\overset{k_{1n_1}}{\rightleftarrows}} & C_{1n_1} \underset{k_{-1(n_1+1)}}{\overset{k_{1(n_1+1)}}{\rightleftarrows}} P_2 + E_1 \\
P_2 + E_2 & \underset{k_{-21}}{\overset{k_{21}}{\rightleftarrows}} & C_{21} & \underset{k_{-22}}{\overset{k_{22}}{\rightleftarrows}} & C_{22} & \underset{k_{-23}}{\overset{k_{23}}{\rightleftarrows}} \quad \cdots \quad \underset{k_{-2n_1}}{\overset{k_{2n_1}}{\rightleftarrows}} & C_{2n_2} \underset{k_{-2(n_2+1)}}{\overset{k_{2(n_2+1)}}{\rightleftarrows}} P_3 + E_2 \\
\vdots & & & & & \vdots & & \vdots \\
P_m + E_m & \underset{k_{-m1}}{\overset{k_{m1}}{\rightleftarrows}} & C_{m1} & \underset{k_{-m2}}{\overset{k_{m2}}{\rightleftarrows}} & C_{m2} & \underset{k_{-m3}}{\overset{k_{m3}}{\rightleftarrows}} \quad \cdots \quad \underset{k_{-mn_m}}{\overset{k_{mn_m}}{\rightleftarrows}} & C_{mn_m} \underset{k_{-m(n_m+1)}}{\overset{k_{m(n_m+1)}}{\rightleftarrows}} P_1 + E_m
\end{array} \tag{22}$$

where $m \in \mathbb{Z}_{>0} \setminus \{1\}$ and $n_1, \ldots, n_m \in \mathbb{Z}^m_{>0}$. We impose the follow restrictions on on rate constants:

$$k_{ij} > 0 \text{ and } k_{-ij} \geq 0 \text{ for } j = 1, \ldots, n_i \text{ and } i = 1, \ldots, m.$$

There are $\beta := 2m + (n_1 + n_2 + \cdots + n_m)$ species in this network.

Remark 3. The notation we use to express the network in (22) is based on the scheme in [15], but with a few changes. In keeping with the notion that a Phosphorylation network has n sites, we use n_i to denote the number of sites in component i, whereas m_i is used in [15]. We use m to represent the number of components in the network.

Note that when $m = 2$ and $n_1 = n_2$ the all-encompassing model reduces to the fully reversible model (10). Hence, by Proposition 3.1, this model captures all of the systems described in Table 1. Additionally, this model also encompasses systems with more than 2 components and different numbers of sites in each component.

This model generalizes the model in [15] by allowing all back-reactions to be optional (irreversibility) and by adding a back-reaction to the final reaction (product inhibition).

Lemma 5.1. *Let Γ be the stoichiometric matrix associated with the reaction network* (22)*. The vector of length β with all 1s (denoted by $\vec{1}$) is an element of ker Γ.*

Proof. Notice that each species in (22) appears in exactly two equations, once as a reactant and once as a product. Hence, the sum of each row of Γ is 0 and $\vec{1} \in \ker \Gamma$. □

By Lemma 5.1, nullity(Γ) $\neq 0$. This implies that nullity(Γ^T) $\neq 0$, giving the following corollary:

Corollary 5.2. *The network* (22) *is conservative. That is, it has a conservation relation.*

In section 6 we establish that the all-encompassing model (22) is globally stable.Unlike our proof that the fully-reversible n-site system (10) is stable in section 4, we do not appeal to the system of ODEs arising from mass-action kinetics. Instead we show appeal to graph-theoretic criteria that hold on the actual chemical reaction network. This supersedes our result in section 4 because the all-encompassing model can be reduced to the fully-reversible model.

6. Proof of Global Stability Using Graph Reduction

In this section we establish that our all-encompassing phosphorylation model (22) is globally stable using graph reductions. First we explain how to construct two types of graphs from a chemical reaction network: an SR-graph and an R-graph. A theorem from [8] (originally from [1]) helps us prove global stability by establishing several graph-theoretic criteria hold on the SR-graph and R-graph. To simplify the argument we use [8] to remove intermediate species to produce a simpler network, on which the same graph-theoretic conditions can be tested to prove stability. Notation is established in sections 6.1-6.3 to be able to apply the results in 6.4. Much of the setup in 6.1-6.4 follows from [8].

Notationally, let $G = (V, E, L)$ be a directed, labeled graph, with vertex set V, edge set E and label set L. Let $X, Y \in E$. A directed edge from X to Y is denoted by $\overrightarrow{XY}$.

6.1. **Assumptions.** To apply the machinery in [8], the following formal assumptions about reaction networks must be satisfied:

(1) $y \to y' \in \mathcal{R} \implies y \leftrightharpoons y', y' \leftrightharpoons y \notin \mathcal{R}$,
(2) $y \leftrightharpoons y' \in \mathcal{R} \implies y' \leftrightharpoons y \notin \mathcal{R}$,
(3) for each $y \in \mathcal{C}$, there exists a reaction in $\mathcal{C}$ that has y as a reactant or product,
(4) for each $i \in \{1, \ldots, n\}$, there exists an $(\alpha_1, \ldots, \alpha_n) \in \mathcal{C}$ such that $a_i > 0$.

Some theorems require the following conditions to hold as well:
(G1) There are no auto-catalytic reactions, meaning that no species can appear as both reactant and product in any reaction.
(G2) Each species in $\mathcal{S}$ takes part in at most two reactions in $\mathcal{R}$.
(G3) The network is conservative, that is, it has a conservation law $c \in \mathbb{R}^n_{>0}$.

Remark 4. In [8], assumptions labeled (r1), (r2) and (r3) are also required by some theorems. According to Remark 1 in [8], these assumptions are satisfied for systems arising from mass-action kinetics (such as our phosphorylation systems), so they are omitted here.

6.2. **Graph Constructions.** Here we explain how to construct two types of graphs derived from a chemical reaction network: directed SR-graphs and R-graphs.

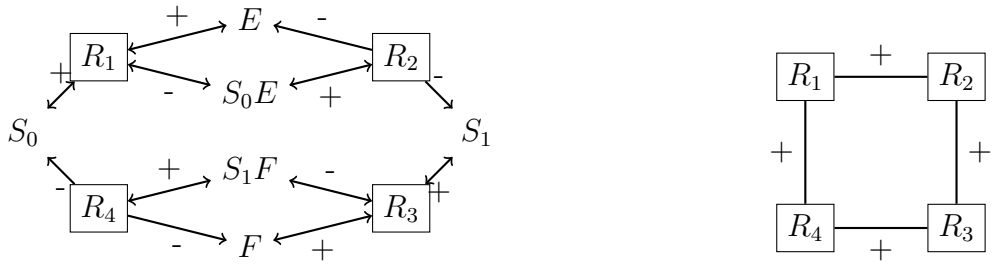
An *SR-graph* (or *directed SR-graph*) is a directed graph constructed from a chemical reaction network. We denote an SR-graph by $G_{SR} = (V_{SR}, E_{SR}, L_{SR})$. The vertex set V_{SR} is the union of all species and reactions in the network (hence the name "SR"). The following rules characterize the edge and label sets:

(1) If a species S is a reactant in any reaction or a product in a reversible reaction, then $\overrightarrow{SR}, \overrightarrow{RS} \in E_{SR}$.
(2) If a species S is a product in an irreversible reaction, then $\overrightarrow{RS} \in E_{SR}$.
(3) Let $SR \in E$. If S is a reactant, $L(S,R) := +$. If S is a product, $L(S,R) := -$. We assign $L(R,S) := L(S,R)$.

An *R-graph* is an undirected graph $G = (V_R, E_R, L_R)$ created from the chemical reaction network. The R-graph can be constructed independently of the SR-graph (which is explained in [8]), but we may abstract it from the SR-graph using the following rules:

(1) The vertex set V_R is the set of reactions in the reaction network.
(2) An edge connects reactions R_i and R_j if there is a length-2 path connecting R_i and R_j in the SR-graph. It is labeled with the opposite of the product of the labels on the length-2 edge. An edge may have more than one label, if there are multiple such paths.

Example 6.1. Consider the 1-site phosphorylation system introduced in Example 2.1. The directed SR-graph and R-graph for this network are shown below.



(A) The SR-graph. (B) The R-graph.

FIGURE 2. The directed SR-graph and R-graph for the 1-site phosphorylation network.

The attributes of the SR-graph and the R-graph that follow help establish the stability of a system (explained in section 6.4).

Definition 6.1. An SR-graph is *R-strongly connected* if there exists a directed path between every pair of reaction vertices.

Definition 6.2. We say an R-graph has the *positive loop property* if every simple loop has an even number of negative edges.

Example 6.2. Again consider the 1-site phosphorylation system introduced in Example 2.1. The SR- and R-graph are shown in Figure 2. The SR-graph is R-strongly connected because it contains the loop

$$R_1 \to S_0E \to R_2 \to S_1 \to R_3 \to S_1F \to R_4 \to S_0 \to R_1.$$

The R-graph vacuously has the positive loop property because it has no negative labels.

From [8], we have the following remark.

Remark 5. Suppose there are p reactions in the system. When the R-graph has the positive loop property, we can define an orthant cone

$$K = \{(x_1, \dots, x_p) \in \mathbb{R}^p \mid \sigma_1 x_1, \dots, \sigma_p x_p \geq 0\} \tag{23}$$

by defining a sign pattern $\sigma = (\sigma_1, \dots, \sigma_p) \in \{\pm 1\}^m$. Let $\sigma_1 := 1$. For $i \in \{2, 3, \dots, p\}$, consider any simple path $1 = i_0 - i_1 - \cdots - i_k = i$ joining 1 and i and set

$$\sigma_i := \prod_{d=1}^{k} L_R(\{R_{i_{d-1}}, R_{i_d}\}). \tag{24}$$

Because the R-graph has the positive loop property, every simple loop has an even number of negative edges, and so any path from 1 to i will result in the same σ_i.

If the R-graph has more than one component, the same procedure is applied to each component, starting by setting $\sigma_1 := 1$ for the smallest index $i \in \{1, \dots, m\}$ such that R_i belongs to that component.

6.3. **Removing Intermediates.** The novelty of [8] is that we can first simplify a chemical reaction network by removing intermediate species before checking conditions on the SR- and R-graphs. This invariance property will be described in section 6.4.

We discuss the required conditions to remove one intermediate. Let $G = (\mathcal{S}, \mathcal{C}, \mathcal{R})$. For each $y = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n_{\geq 0}$ we define the *support of a complex* y by

$$\text{supp } y := \{S_i \in \mathcal{S} \mid \alpha_i > 0\}.$$

To remove an intermediate Y, the following two conditions must be met:

- (l1) Y consists of exactly one species and does not appear in any other complex in the network
- (l2) there exist unique complexes $y = \alpha_1 S_1 + \cdots + \alpha_n S_n$ and $y' = \alpha'_1 S_1 + \cdots + \alpha'_n S_n$ such that
  - (a) either $y \to Y$ or $y \leftrightharpoons Y$ is a reaction
  - (b) either $Y \to y'$ or $Y \leftrightharpoons y'$ is a reaction
  - (c) If $\mathcal{E} := \text{supp } y \cap \text{supp } y'$, then $\displaystyle\sum_{S_i \in \mathcal{E}} \alpha_i S_i = \sum_{S_i \in \mathcal{E}} \alpha'_i S_i =: e$
  - (d) $y - e \to y' - e$, $y' - e \to y - e$, $y - e \leftrightharpoons y' - e$ and $y' - e \leftrightharpoons y - e$ are not reactions in $\mathcal{R}$.

Definition 6.3. Given a network $G = (\mathcal{S}, \mathcal{C}, \mathcal{R})$ and an intermediate Y that satisfies (l1) and (l2), the reduced reaction network $G = (\mathcal{S}^*, \mathcal{C}^*, \mathcal{R}^*)$ is obtained by **removing the intermediate** Y. We define $\mathcal{R}^* := \mathcal{R}^*_c \cup \mathcal{R}^*_Y$, where $\mathcal{R}^*_Y$ is the subset of reactions in $\mathcal{R}$ that do not have Y as a product or reactant and

$$R^*_Y := \begin{cases} \{y - e \leftrightharpoons y' - e\}, \text{ if } y \leftrightharpoons Y, Y \leftrightharpoons y' \in \mathcal{R} \\ \{y - e \to y' - e\}, \text{ if } y \to Y \in \mathcal{R}, \text{or } Y \to y' \in \mathcal{R} \end{cases}. \tag{25}$$

This procedure removes one intermediate. Any number of intermediates may be removed successively if the conditions above are met at each step.

Example 6.3. As an example, consider the 1-site phosphorylation network (2.1). Taking $S_0 + K$ and $S_1 + K$ to be the unique complexes y and y' required by (l2), we can remove the

intermediate S_0K, producing the reduced network

$$\begin{aligned} S_0 &\longrightarrow S_1 \\ S_1 + F &\rightleftharpoons S_1F \longrightarrow S_0 + F \end{aligned}$$

Notice that K is also removed because it is in both $S_0 + K$ and $S_1 + K$. If we try to remove S_1F we must take $S_1 + F$ and $S_0 + F$ to be y and y'. Then $e = F$, but $y' - e \to y - e$ is a reaction, which violates (l2). Thus, we cannot remove any other intermediates.

We use successive removal of intermediates to simplify the all-encompassing model (10) in the lemma below. Notationally, let $\rightleftharpoons$ denote a reaction that may or may not be reversible.

Lemma 6.1. *The following network can be obtained from the all-encompassing model* (10) *using successive removal of intermediates:*

$$\begin{aligned} R_1^*: &\qquad P_1 \rightleftharpoons P_m \\ R_2^*: &\qquad P_m + E_m \rightleftharpoons C_{mn_m} \\ R_3^*: &\qquad C_{mn_m} \rightleftharpoons P_1 + E_m \end{aligned} \tag{26}$$

Proof. First we claim that we can remove C_{ij} for $i = 1, \dots, m$ and $j = 1, \dots, n_i$. Consider the component with index $i \in \{1, \dots, m\}$. We show inductively that we can remove the intermediates $C_{i1}, C_{i2}, \dots, C_{in_i}$. Notice that these complexes contain one species and do not appear elsewhere in the system, so assumption (I1) is satisfied.

This component contains the reactions $P_i + E_i \rightleftharpoons C_{i1}$ and $C_{i1} \rightleftharpoons C_{i2}$. The complexes $P_1 + E_1$ and C_{i2} do not share any species, and they satisfy (I2). Hence, we can remove the intermediate C_{i1}. This establishes the base case.

Suppose that we have removed intermediates $C_{i1}, C_{i2}, \dots, C_{i(k-1)}$, where $1 \leq k \leq n_i$. If $k \neq n_i$, component i contains the reactions $P_i + E_i \rightleftharpoons C_{ik}$ and $C_{ik} \rightleftharpoons C_{i(k+1)}$. Similarly to the base case, the complexes $P_i + E_i$ and $C_{i(k+1)}$ satisfy (I2), so we can remove the intermediate C_{ik}. In the case that $k = n_i$, the complexes $P_i + E_i$ and $P_{i+1} + E_i$ satisfy (I2) as well and so C_{in_i} can be removed. The neighboring complexes share E_i and so it is also removed as per (25).

Successively remove all C_{ij} in each component, except C_{mn_m}, results in the network

$$\begin{aligned} P_1 &\rightleftharpoons P_2 \\ P_2 &\rightleftharpoons P_3 \\ &\vdots \qquad\qquad \vdots \\ P_{m-1} &\rightleftharpoons P_m \\ P_m + E_m &\rightleftharpoons C_{mn_m} \rightleftharpoons P_1 + E_m \end{aligned} \tag{27}$$

If $m = 2$, we are done. Otherwise, we claim that we can successively remove $P_2, P_3, \dots, P_{m-1}$. Notice that (l1) is satisfied because each P_i is contained in just one complex. We can remove P_2 because P_1 and P_3 satisfy (l2). If we have removed $P_2, \dots, P_{k-1}$, where $k \leq m - 1$, then P_1 and P_{k+1} satisfy (l2). This results in the network (26).

□

Remark 6. We are able to remove any further intermediates from the reduced network 26 produced by Lemma 6.1. However, there are cases when simpler reductions from the all-encompassing model (22) are possible. Consider the case that $m \geq 3$. Using arguments similar to those in Lemma 6.1 we could remove all C_{ij}, including C_{mm_n}, and $P_2, \ldots, P_{m-2}$ (if $m \geq 4$) to produce the slightly simpler network

$$P_1 \rightleftharpoons P_{m-1}$$
$$P_{m-1} \rightleftharpoons P_m$$
$$P_m \rightleftharpoons P_1$$

No further reductions can be made without violating (l2d).

However, the reduced network (26) is minimal in the $m = 2$ case. For example, if we try to remove C_{2n_2}, we would be forced to take the neighbors $P_2 + E_2$ and $P_1 + E_2$ as y and y' in accordance with (l2). However, $P_1 \rightleftharpoons P_2$ is already a reaction, which violates (l2d).

6.4. **Stability Results.** Now we can state the results from [8] necessary for our proof. The following Lemma (Theorems 1 and 2 in [7]) establishes that the desired properties on the SR- and R-graphs are invariant under successive removal of intermediates.

Lemma 6.2. *Let G be a reaction network satisfying (G1) and (G2). Suppose G^* is a reaction network obtained from G by successive removal of intermediates. Then G also satisfies (G1) and (G2) and*

(1) the directed SR-graph of G^* is R-strongly connected if, and only if, the directed SR-graph of G is strongly connected.
(2) The R-graph of G^* has the positive loop property if, and only if, the R-graph of G has the positive loop property.

Furthermore, properties involving the stiochimetric matrix and the cone from Remark 5 are invariant under successive removal of intermediates.

Lemma 6.3. *Let G be a reaction network satisfying (G1) and (G2). Suppose G^* is a reaction network obtained from G by the successive removal of intermediates. Suppose, in addition, that the R-graph of G^* has the positive loop property and the SR-graph of G^* is R-strongly connected. Let Γ and Γ^* be the stoichiometric matrices of G and G^*, and K and K^* the orthant cones constructed in Remark 5 from the R-graphs of G and G^*, respectively. Then,*

$$\ker \Gamma \cap K = \{0\} \iff \ker \Gamma^* \cap K^* = \{0\},$$

and

$$\ker \Gamma \cap \operatorname{int} K \neq \emptyset \iff \ker \Gamma^* \cap \operatorname{int} K^* \neq \emptyset.$$

To state the main stability result we require the following definitions from [7].

Definition 6.4. The **ω-limit set** is

$$\omega(s_0) := \bigcap_{\tau \gg 0} \bigcup_{t \gg \tau} \{\sigma(t, s_0)\},$$

where $\sigma(t, s_0)$ is the trajectory of $x(t)$ starting from initial condition s_0.

Definition 6.5. The flow of (6) is said to be **bounded-persistent** if $\omega(s_0) \cap \partial\mathbb{R}^n_{\geq 0} = \emptyset$ for each $s_0 \in \mathbb{R}^n_{>0}$.

The following lemma (Propositions 2 and 3 in [8]) provides sufficient conditions for both local and global stability.

Lemma 6.4. *Let G be a reaction network satisfying (G1)-(G3) and (r1)-(r3). Suppose that the flow of* (6) *is bounded-persistent. Suppose, in addition, that the R-graph of G has the positive loop property and that the directed SR-graph of G is strongly connected. Then, either,*

(1) $\ker \Gamma \cap K = \{0\}$ and there exists a Lebesgue measure-zero set $D \subseteq \mathbb{R}^n_{>0}$ such that all solutions to (6) *starting in $\mathbb{R}^n_{>0} \setminus D$ converge to the set of equilibria, or*
(2) $\ker \Gamma \cap int\ K \neq \emptyset$ and all solutions of (6) *starting in $\mathbb{R}^n_{>0}$ converge to an equilibrium. Furthermore, this equilibrium is unique within each stoichometric compatibility class.*

Appendix A shows how bounded-persistence can be established with graph-theoretic criteria from [1]. Hence, each condition required by Lemma 6.4 can be shown using graph-theoretic criteria.

6.5. **Proof of Global Stability of All-Encompassing Model.**

Theorem 6.5. *For any chemical reaction system* (6) *arising from all-encompassing network* (22) *and any choice of rate constants,*

(1) each stoichiometric compatibility class $\mathcal{P}$ contains a unique steady state η,
(2) η is a positive steady state, and
(3) η is the global attractor of $\mathcal{P}$.

Proof. By inspection, the full network (22) satisfies assumptions (G1) and (G2). Assumption (G3) is satisfied by Corollary 5.2 and assumptions (r1)-(r3) are satisfied by Remark 4.

By Lemmas 6.2, 6.3 and 6.4, it suffices to show that each of the following properties is satisfied by the reduced network (26) or the original network (22):

(1) the network bounded persistent,
(2) $\ker \Gamma \cap \text{int } K \neq \emptyset$,
(3) the SR-graph is R-strongly connected, and
(4) the R graph has the positive loop property.

By Lemma (A.2), the reduced network is bounded-persistent and so property 1 holds.

Applying Lemma 5.1, for the full network (22), we have $\vec{1} \in \ker \Gamma \cap \text{int } K$ and so property 2 holds.

The SR- and R-graphs resulting from the reduced network (26) are shown in Figure 3. A dashed edge is present if and only if the corresponding component in the original network (22) is fully-reversible.

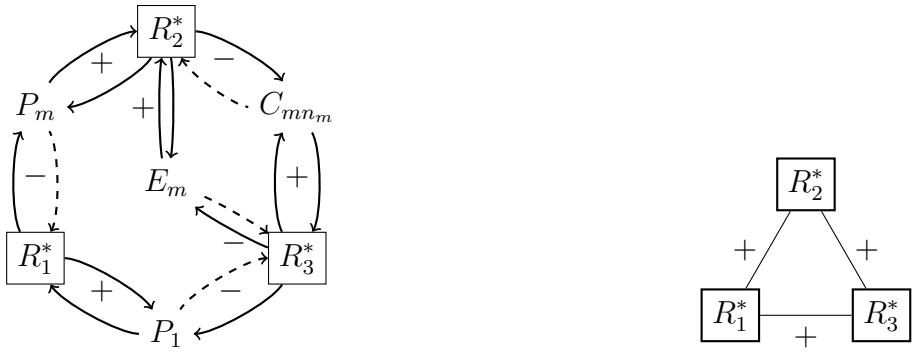
Regardless of our choice of rate constants, the SR-graph always contains the loop

$$R_1^* \to P_m \to R_2^* \to C_{mn_m} \to R_3^* \to P_1 \to R_1^*.$$

Hence, there exists a directed path between any two reactions and, by definition, the SR-graph is R-strongly connected.

Notice that each length-two path connecting a pair of reactions in the SR-graphs has edges with opposite signs. It follows that each edge in the R-graph has a positive label. Thus, the R-graph has no negative labels and so it vacuously has the positive loop property.

□



(A) directed SR-graph (B) R-graph.

FIGURE 3. The SR-graph and R-graph of the reduced system (26).

7. Remarks

Here we discuss how our phosphorylation networks compare to other types of phosphorylation systems in the literature.

Remark 7. Theorem 6.5 generalizes the stability result established by Rao in [15] by allowing the final reaction in each component to be reversible and allowing all other reactions to be either reversible or irreversible. In [15], Rao built a Lyapunov function to prove global stability. It would be interesting to see if Rao's proof could easily be extended as another means to establish Theorem 6.5.

Remark 8. Only recently have there been studies of mixed phosphorylation mechanisms (partially distributive, partially processive) [3]. Suwanmajo and Krishnan proved that such a network, in which phosphorylation is distributive and dephosphorylation is processive (or, by symmetry, vice-versa), is *not* multistationary [17]. Thus, it always admits a unique steady state, via a standard application of the Brouwer fixed-point theorem. This proves half of a conjecture that Conradi and Shiu posed [6]. Perhaps surprisingly, the other half of the conjecture was disproven: in contrast with processive systems (§1.1), mixed systems need not be globally stable: they can be oscillatory [17]!

Remark 9. There are examples of phosphorylation networks in the literature that have more reactions that those in our all-encompassing model (22). In [11], Gunawardena proposes a 2-site processive phosphorylation network in which, in addition to the usual edges, ES_0 reacts to form $E + S_2$:

$$E + S_0 \rightleftharpoons ES_0 \begin{array}{l} \nearrow E + S_2 \\ \searrow E + S_1 \end{array}$$

Unfortunately our proof of Theorem 4.2 cannot be extended to establish the stability of such networks. Suppose we add an additional reaction edge to the all-encompassing model (10). It is easy to show that the new reaction vector is a linear combination of other columns in the stoichiometric matrix Γ. Hence the rank of Γ is invariant under adding new reactions but no new complexes. However, the number of reactions m increases and so the first condition of Lemma 4.1 is not met. Thus, our argument does not apply if we add a new reaction but no new complexes.

Additionally, our proof of Theorem 6.5, establishing the stability of the all-encompassing model, cannot be extended in this way because condition (G2) does not allow a complex to be present in more than two reactions.

Remark 10. We showed that the generalized all-encompassing model is globally stable regardless of whether any of the reactions are reversible or irreversible. If a reversible system is stable using the graph-theoretic methods in section 6, is it also true that the associated irreversible system is also stable? We conjecture that this is not true in general.

Remark 11. The graph reduction tools in [8] used to establish global stability of the reduced all-encompassing model in section 6 do not directly consider dynamical systems arising from mass-action kinetics. However, the graph-theoretic conditions used are equivalent to some of those required by [2], which was used to establish stability of the fully-reversible model in section 4.

More specifically, recall the translated version of the fully-reversible system (14). The work in [1] establishes that

(1) The R-graph is R-strongly connected if and only if the dynamical system (14) is monotone,
(2) if condition (1) holds, then the SR-graph is R-strongly connected if and only if the Jacobian matrix is irreducible along trajectories, and
(3) condition (2) implies the dynamical system (14) is strongly monotone.

This means that the graph conditions described in section 6 could be used to establish that the fully-reversible system from section 3 is strongly monotone (a condition of 4.1). The proof would be similar to Example 4 in [8].

Acknowledgments. This work was funded by NSF grant DMS-1460766. The author would like to thank Dr. Anne Shiu for her guidance at the 2016 Texas A&M University Mathematics REU. Some parts of this paper are adapted from [6] with permission.

A. Bounded-Persistence

Here we provide the background necessary to understand bounded persistence. Intuitively, a system is bounded persistent if no trajectory vanishes asymptotically as $t \to \infty$. This idea is formalized by Definition 6.5.

A.1. **Definitions.** Using [7], bounded persistence can be established through the use of P-semiflows and siphons, defined below.

Definition A.1. A **P-semiflow** of a reaction network is any nonzero vector $v \in \mathbb{R}^n_{>}0$ such that $\Gamma^T v = 0$, i.e. v is a positive conservation law.

Definition A.2. A nonempty subset of species $\Sigma \subseteq S$ is called a **siphon** if every reaction which has a product in Σ also has a reactant in Σ. A siphon is said to be minimal if it does not properly contain any other siphon.

Definition A.3. A reaction network is said to have the **siphon/P-semiflow property**, or satisfy the siphon/P-semiflow condition, if every siphon contains the support of a P-semiflow.

This theorem from [7] states that the siphon/P-semiflow property is a sufficient condition to establish bounded-persistence.

Theorem A.1. *If a reaction network has the siphon/P-semiflow property, then it is bounded-persistent.*

A.2. **Proof for the Reduced All-Encompassing Model.**

Lemma A.2. *The reduced network* (26) *is bounded-persistent.*

Proof. Assume the ordering of species P_1, P_m, C_{mn_m}, E_m. The reduced network (26) has the stoichiometric matrix

$$\Gamma^* = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

The minimal siphons of the reduced network (26) are $\Sigma_1 := \{P_1, P_1, C_{mn_m}\}$ and $\Sigma_2 := \{P_1, C_{mn_m}, E_m\}$. By inspection, $(0,0,1,1), (1,1,1,0) \in \ker\left(\Gamma^*\right)^T$ and so $(0,0,1,1)$ and $(1,1,1,0)$ are P-semiflows. The siphon Σ_1 contains the support of $(1,1,1,0)$ and Σ_2 contains the support of $(0,0,1,1)$. Thus, by definition, the reduced network (26) has the siphon/P-semiflow property. It follows from Theorem A.1 that the reduced network is bounded persistent. □

References

[1] D. Angeli, P. De Leenheer, and E. Sontag. Graph-theoretic characterizations of monotonicity of chemical networks in reaction coordinates. *Journal of Mathematical Biology*, 61(4):581–616, 2010.

[2] D. Angeli and E. D. Sontag. Translation-invariant monotone systems, and a global convergence result for enzymatic futile cycles. *Nonlinear Anal. Real World Appl.*, 9(1):128–140, 2008.

[3] K. Aoki, M. Yamada, K. Kunida, S. Yasuda, and M. Matsuda. Processive phosphorylation of ERK MAP kinase in mammalian cells. *P. Natl. Acad. Sci. USA*, 108(31):12675–12680, 2011.

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

[5] M. Banaji and J. Mierczyński. Global convergence in systems of differential equations arising from chemical reaction networks. *J. Differential Equations*, 254(3):1359–1374, 2013.

[6] C. Conradi and A. Shiu. A global convergence result for processive multisite phosphorylation systems. *Bulletin of Mathematical Biology*, 77(1):126–155, 2015.

[7] M. M. de Freitas, E. Feliu, and C. Wiuf. Intermediates, catalysts, persistence, and boundary steady states. *arXiv preprint arXiv:1509.06034*, 2015.

[8] M. M. de Freitas, C. Wiuf, and E. Feliu. Intermediates and generic convergence to equilibria. *arXiv preprint arXiv:1606.09480*, 2016.

[9] H. Errami, M. Eiswirth, D. Grigoriev, W. M. Seiler, T. Sturm, and A. Weber. Detection of Hopf bifurcations in chemical reaction networks using convex coordinates. *J. Comput. Phys.*, 291:279–302, 2015.

[10] J. Gunawardena. Multisite protein phosphorylation makes a good threshold but can be a poor switch. *PNAS*, 102(41):14617–14622, 2005.

[11] J. Gunawardena. Distributivity and processivity in multisite phosphorylation can be distinguished through steady-state invariants. *Biophys. J.*, 93(11):3828–3834, 2007.

[12] J. Hell and A. D. Rendall. A proof of bistability for the dual futile cycle. *Nonlinear Anal.-Real*, 24:175–189, 2015.

[13] N. I. Markevich, J. B. Hoek, and B. N. Kholodenko. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *The Journal of Cell Biology*, 164:353 – 359, 2004.

[14] P. Patwardhan and W. T. Miller. Processive phosphorylation: Mechanism and biological importance. *Cellular Signalling*, 19(11):2218–2226, 2007.

[15] S. Rao. Global stability of a class of futile cycles. *Journal of Mathematical Biology*, pages 1–18, 2016.

[16] H. L. Smith. *Monotone dynamical systems: An introduction to the theory of competitive and cooperative systems*, volume 41 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1995.

[17] T. Suwanmajo and J. Krishnan. Mixed mechanisms of multi-site phosphorylation. 2015.

Department of Mathematical Sciences, Ripon College, Ripon, WI 54971
E-mail address: eithunm@ripon.edu