



Conditions for extinction events in chemical reaction networks with discrete state spaces

Matthew D. Johnston¹ · David F. Anderson² ·
Gheorghe Craciun^{2,3} · Robert Brijder⁴

Received: 15 March 2017 / Revised: 7 August 2017 / Published online: 26 September 2017
© Springer-Verlag GmbH Germany 2017

Abstract We study chemical reaction networks with discrete state spaces and present sufficient conditions on the structure of the network that guarantee the system exhibits an extinction event. The conditions we derive involve creating a modified chemical reaction network called a domination-expanded reaction network and then checking properties of this network. Unlike previous results, our analysis allows algorithmic implementation via systems of equalities and inequalities and suggests sequences of reactions which may lead to extinction events. We apply the results to several networks including an EnvZ-OmpR signaling pathway in *Escherichia coli*.

Keywords Reaction network · Reaction graph · Extinction · Stochastic process · Petri net

Mathematics Subject Classification 92C42 · 60J27

1 Introduction

Continuous state differential equations are a popular modeling choice for the chemical concentrations of biochemical reaction networks in several disciplines, including

✉ Matthew D. Johnston
matthew.johnston@sjsu.edu

¹ Department of Mathematics, San José State University, San Jose, CA 95192, USA

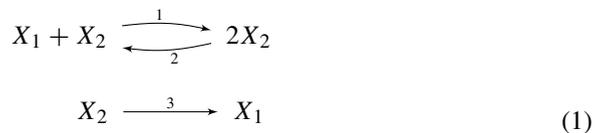
² Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706, USA

³ Department of Biomolecular Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA

⁴ Department WET-INF, Hasselt University, Diepenbeek, Belgium

industrial chemistry and systems biology. However, differential equations should only be used to model chemical concentrations when the counts of the reactant species are high (Kurtz 1972; Anderson and Kurtz 2011, 2015). When the multiplicities of some of the individual species are low, as is often the case in enzymatic and genetic systems, it is important to use a model with a discrete state space which tracks individual molecular counts.

Predictions pertaining to the long-term behavior of a particular system can change dramatically depending upon whether the system is modeled with a continuous or discrete state space. In particular, discrete space models may exhibit an extinction event where none exists in the corresponding continuous state model. For example, consider the following chemical reaction network:



where the labels correspond to the enumeration of the reactions. The deterministic mass action model predicts an asymptotically stable steady state for a wide range of parameter values. However, for the discrete space model with stochastic mass-action kinetics and $M = X_1^\# + X_2^\#$, where $X_i^\#$ is the count of species X_i , the state $\{X_1^\# = M, X_2^\# = 0\}$ is the inevitable absorbing state regardless of parameter values. This extinction event can be achieved by reaction 3 occurring until the count of species X_2 is zero, at which point no further reactions may occur.

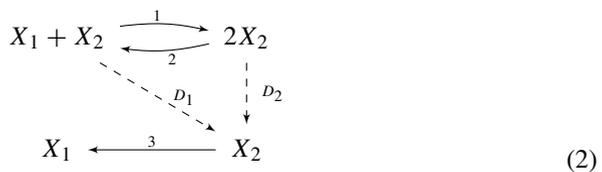
Several frameworks exist for tracking trajectories of discrete state chemical reaction systems, including those of continuous time Markov chains (Anderson and Kurtz 2011, 2015) and stochastic Petri nets (Bause and Kritzinger 2002). In these settings, the admissible transitions between states are assumed to occur randomly at a known rate and the occurrence of each reaction instantaneously updates the system according to the stoichiometry of the associated reaction. Analysis of such systems is typically conducted by generating sample trajectories [through a stochastic simulation algorithm, e.g. Gillespie's Algorithm (Gillespie 1976) or the next reaction method (Gibson and Bruck 2000; Anderson 2007)], by analyzing the evolution of the probability distribution via Kolmogorov's forward equations (i.e. the *chemical master equation*), by characterizing the stationary distributions of the models (Anderson et al. 2011), or by studying the stochastic equations for the model (Anderson and Kurtz 2011, 2015).

The study of extinction events in discrete interaction models is well-established in population dynamics and epidemic modeling, but the corresponding study in systems biology has only recently gained widespread attention. Anderson et al. (2014) described a large class of systems for which an extinction event necessarily occurs in the discrete model. Interestingly, this class of models had previously been shown by Shinar and Feinberg (2010) to have a particular "robustness" when modeled with deterministic ordinary differential equations. Brijder (2015) utilized tools from Petri Net Theory to further extend the scope of networks known to have extinction behavior by relating a kernel condition introduced in Anderson et al. (2014) to the T -invariants

of the corresponding Petri net. Related recent work analyzing transient and post-extinction behavior in discrete chemical reaction systems has been conducted by Enciso (2016) and Anderson et al. (2017).

In this paper, we further develop a network-based approach to determining when discrete-space chemical reaction systems may exhibit an extinction event. Our main results, Lemma 2 and Theorem 1, state that a chemical reaction network with a discrete state space exhibits an extinction event if there is a modified network, called the *domination-expanded reaction network*, on which a particular set of inequalities on the edges *cannot* be satisfied. The conditions we present may be summarized as systems of equalities and inequalities and, like Corollary 2 of Brijder (2015), suggest computational implementation.

For example, the network (1) can be correspond with the following domination-expanded network



where we treat the arrows labeled D_1 and D_2 as reactions in a new network. As will be described later in the paper, the network (2) can be furthermore corresponded to the following system of equalities and inequalities

$$\left\{ \begin{array}{l}
 \text{(Cond. 1):} \quad (\alpha_R)_1 = 0, (\alpha_D)_2 = 0 \\
 \text{(Cond. 2):} \quad -(\alpha_R)_1 + (\alpha_R)_2 + (\alpha_R)_3 = 0 \\
 \quad \quad \quad (\alpha_R)_1 - (\alpha_R)_2 - (\alpha_R)_3 = 0 \\
 \text{(Cond. 3):} \quad (\alpha_R)_3 \geq (\alpha_D)_1 \geq (\alpha_R)_2 \geq 0.
 \end{array} \right. \tag{3}$$

on the vector of reaction counts $\alpha = ((\alpha_R)_1, (\alpha_R)_2, (\alpha_R)_3, (\alpha_D)_1, (\alpha_D)_2) \in \mathbb{Z}_{\geq 0}^5$. Since there is no vector $\alpha \neq \mathbf{0}$ satisfying (3), Theorem 1 guarantees that the chemical reaction network (1) on a discrete state space has an extinction event. A computational implementation of Theorem 1 is explored in the companion paper of Johnston (2017).

The notation of the paper is drawn from *chemical reaction network theory*, which has proven effective for relating topological properties of a network’s reaction graph to its admissible qualitative dynamical behaviors (Horn 1972; Horn and Jackson 1972; Feinberg 1972, 1987, 1988). The notions introduced here may be equivalently defined in the context of Petri nets (Bause and Kritzinger 2002; Brijder 2015). We also adopt the following common notation throughout the paper:

- $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} \mid x \geq 0\}$ and $\mathbb{R}_{> 0} = \{x \in \mathbb{R} \mid x > 0\}$,
- for $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}_{\geq 0}^n$, we define $\text{supp}(\mathbf{v}) = \{i \in \{1, \dots, n\} \mid v_i > 0\}$,
- for a set $X = \{X_1, X_2, \dots, X_n\}$ of indexed elements and a subset $W \subseteq X$, we define $\text{supp}(W) = \{i \in \{1, \dots, n\} \mid X_i \in W\}$,
- for a subset $W \subseteq X$, we define the complement $W^c = \{x \in X \mid x \notin W\}$,
- for $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, we define $\mathbf{v} \leq \mathbf{w}$ if $v_i \leq w_i$ for each $i \in \{1, \dots, n\}$.

2 Background

We outline the background notation and terminology relevant to the study of *chemical reaction network theory* (CRNT). (For further background, see Martin Feinberg's online lecture notes [Feinberg 1979](#).)

2.1 Chemical reaction networks

The fundamental object of interest in CRNT is the following.

Definition 1 A *chemical reaction network* (CRN) is given by a triple of finite sets $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ where:

1. The *species set* $\mathcal{S} = \{X_1, \dots, X_m\}$ contains the species of the CRN.
2. The *reaction set* $\mathcal{R} = \{R_1, \dots, R_r\}$ consists of ordered pairs $(y, y') \in \mathcal{R}$ where

$$y = \sum_{i=1}^m y_i X_i \quad \text{and} \quad y' = \sum_{i=1}^m y'_i X_i, \quad (4)$$

and where the values $y_i, y'_i \in \mathbb{Z}_{\geq 0}$ are the *stoichiometric coefficients*. We will also write reactions (y, y') as $y \rightarrow y'$.

3. The *complex set* \mathcal{C} consists of the linear combinations of the species in (4). Specifically, $\mathcal{C} = \{y \mid y \rightarrow y' \in \mathcal{R}\} \cup \{y' \mid y \rightarrow y' \in \mathcal{R}\}$. The number of distinct complexes is denoted $|\mathcal{C}| = n$.

Allowing for a slight abuse of notation, we will let y denote both the complex itself and the complex vector $y = (y_1, \dots, y_m)^T \in \mathbb{Z}_{\geq 0}^m$.

We assume an arbitrary but fixed ordering of the species, reactions and complexes. It is common to impose that a CRN does not contain any self-loops (i.e. reactions of the form $y \rightarrow y$). Since this assumption is not used in our results, and since it is common to allow self-loops in Petri Net Theory, we will not make this assumption here.

The interpretation of reactions as directed edges naturally gives rise to a reaction graph $G = (V, E)$ where the set of vertices is given by the complexes (i.e. $V = \mathcal{C}$) and the set of edges is given by the reactions (i.e. $E = \mathcal{R}$). The following terminology will be used.

- (i) A complex y is *connected* to a complex y' if there exists a sequence of complexes $y = y_{\mu(1)}, y_{\mu(2)}, \dots, y_{\mu(\ell)} = y'$ such that either $y_{\mu(k)} \rightarrow y_{\mu(k+1)}$ or $y_{\mu(k+1)} \rightarrow y_{\mu(k)}$ for all $k \in \{1, \dots, \ell - 1\}$.
- (ii) There is a *path* from y to y' if there is a sequence of distinct complexes such that $y = y_{\mu(1)} \rightarrow y_{\mu(2)} \rightarrow \dots \rightarrow y_{\mu(\ell)} = y'$.
- (iii) A maximal set of mutually connected complexes is called a *linkage class* (LC) while a maximal set of mutually path-connected complexes is called a *strong linkage class* (SLC). The set of LCs will be denoted \mathcal{L} while the set of SLCs will be denoted \mathcal{W} .
- (iv) An SLC $W \in \mathcal{W}$ is called *terminal* if there are no outgoing reactions, i.e. $y \in W$ and $y \rightarrow y' \in \mathcal{R}$ implies $y' \in W$. The set of terminal SLCs will be denoted

- $\mathcal{T} \subseteq \mathcal{W}$. A complex $y \in \mathcal{C}$ is called terminal if it belongs to a terminal SLC, and a reaction $y \rightarrow y' \in \mathcal{R}$ is terminal if y is terminal.
- (v) A set $\mathcal{Y} \subseteq \mathcal{C}$ is called an *absorbing complex set* if it contains every terminal complex and has no outgoing edges, i.e. $y \in \mathcal{Y}$ and $y \rightarrow y' \in \mathcal{R}$ implies $y' \in \mathcal{Y}$. A complex $y \in \mathcal{Y}$ is called \mathcal{Y} -*interior*, and a reaction $y \rightarrow y' \in \mathcal{R}$ is called \mathcal{Y} -*interior* if y is \mathcal{Y} -interior; otherwise they are \mathcal{Y} -*exterior*.

Absorbing complex sets are a generalization of the set of terminal complexes of a CRN, since they must contain, but may be strictly larger than, this set. Note that the set of terminal complexes is an absorbing complex set of the CRN, as is the set $\mathcal{Y} = \mathcal{C}$. We will be particularly interested in the case where \mathcal{Y} is the set of terminal complexes, as this provides the foundation upon which our main results are built.

To each reaction $y \rightarrow y' \in \mathcal{R}$ we associate a *reaction vector* $y' - y \in \mathbb{Z}^m$ which tracks the net gain and loss of each chemical species as a result of the occurrence of this reaction. The *stoichiometric subspace* is defined by

$$S = \text{span} \{y' - y \in \mathbb{Z}^m \mid y \rightarrow y' \in \mathcal{R}\}.$$

The *stoichiometric matrix* $\Gamma \in \mathbb{Z}^{m \times r}$ is the matrix with the reaction vectors as columns. A CRN is said to be *conservative* (respectively, *subconservative*) if there exists a $\mathbf{c} \in \mathbb{Z}_{>0}^m$ such that $\mathbf{c}^T \Gamma = \mathbf{0}^T$ (respectively, $\mathbf{c}^T \Gamma \leq \mathbf{0}^T$). The vector \mathbf{c} is called a *conservation vector* (respectively, *subconservation vector*).

We present three examples in order to illustrate the definitions.

Example 1 Reconsider the CRN given by (1) in the introduction. This CRN has the sets $\mathcal{S} = \{X_1, X_2\}$, $\mathcal{R} = \{X_1 + X_2 \rightarrow 2X_2, 2X_2 \rightarrow X_1 + X_2, X_2 \rightarrow X_1\}$, and $\mathcal{C} = \{X_1 + X_2, 2X_2, X_2, X_1\}$. The linkage classes are

$$\mathcal{L} = \{\{X_1 + X_2, 2X_2\}, \{X_2, X_1\}\}$$

while the SLCs are

$$\mathcal{W} = \{\{X_1 + X_2, 2X_2\}, \{X_2\}, \{X_1\}\}.$$

Note that SLCs may consist of singletons. The terminal SLCs are

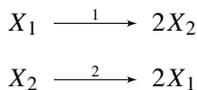
$$\mathcal{T} = \{\{X_1 + X_2, 2X_2\}, \{X_1\}\}.$$

The stoichiometric matrix is as follows:

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

The stoichiometric subspace is given by $S = \text{span}\{(1, -1)^T\}$, and there is the conservation vector $\mathbf{c} = (1, 1)^T$.

Example 2 Consider the following CRN:

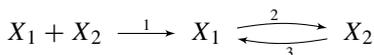


The set of terminal complexes is $\{2X_2, 2X_1\}$. There are several additional choices for absorbing complex sets, including $\mathcal{Y} = \{X_1, 2X_2, 2X_1\}$ and $\mathcal{Y} = \{2X_2, X_2, 2X_1\}$. The stoichiometric matrix is as follows:

$$\Gamma = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}.$$

The stoichiometric subspace is given by $S = \text{span}\{(-1, 2)^T, (2, -1)^T\} = \mathbb{R}^2$. There is no vector $\mathbf{c} \in \mathbb{Z}_{>0}^2$ for which $\mathbf{c}^T \Gamma \leq \mathbf{0}^T$, so the CRN is not conservative or subconservative.

Example 3 Consider the following CRN:



The stoichiometric matrix is as follows:

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

There is no vector $\mathbf{c} \in \mathbb{R}_{>0}^2$ such that $\mathbf{c}^T \Gamma = \mathbf{0}^T$, so the CRN is not conservative; however, the vector $\mathbf{c} = (1, 1)^T$ has the property that $\mathbf{c}^T \Gamma = (-1, 0, 0) \leq \mathbf{0}$ so that the CRN is subconservative.

2.2 Chemical reaction networks with discrete state spaces

For CRNs with discrete state spaces, we let $\mathbf{X} = (X_1^\#, \dots, X_m^\#)^T \in \mathbb{Z}_{\geq 0}^m$ denote a *discrete state* where $X_i^\#$ corresponds to the molecular count of species X_i . For brevity, we often call a discrete state simply a *state*. We will say that a complex $y \in \mathcal{C}$ is *charged* at state $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$ if $X_i^\# \geq y_i$ for all $i \in \{1, \dots, m\}$. Note that a reaction may only occur from a state \mathbf{X} if the species counts are sufficient for the source complex of that reaction.

We introduce the following terminology, which is adapted from the conventions of stochastic processes.

Definition 2 Consider a CRN on a discrete state space. Then:

1. A state $\mathbf{X} \in \mathbb{Z}_{>0}^m$ *reacts to* a state $\mathbf{Y} \in \mathbb{Z}_{\geq 0}^m$ (denoted $\mathbf{X} \rightarrow \mathbf{Y}$) if there is a reaction $y \rightarrow y' \in \mathcal{R}$ such that $\mathbf{Y} = \mathbf{X} + y' - y$ and y is charged at state \mathbf{X} .
2. A state $\mathbf{Y} \in \mathbb{Z}_{\geq 0}^m$ is *reachable* from a state $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$ (denoted $\mathbf{X} \rightsquigarrow \mathbf{Y}$) if there exists a sequence of states such that $\mathbf{X} = \mathbf{X}_{\nu(1)} \rightarrow \mathbf{X}_{\nu(2)} \rightarrow \dots \rightarrow \mathbf{X}_{\nu(l)} = \mathbf{Y}$.
3. A state $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$ is *recurrent* if, for any $\mathbf{Y} \in \mathbb{Z}_{\geq 0}^m$, $\mathbf{X} \rightsquigarrow \mathbf{Y}$ implies $\mathbf{Y} \rightsquigarrow \mathbf{X}$; otherwise, the state is *transient*.

We let $\mathbf{X}(t) = (X_1^\#(t), \dots, X_m^\#(t))^T \in \mathbb{Z}_{\geq 0}^m$ denote the discrete state of our system at time t , so that the system evolves as follows:

$$\mathbf{X}(t) = \mathbf{X}(0) + \Gamma \mathbf{N}(t) \tag{5}$$

where $\mathbf{N}(t) = (N_1(t), \dots, N_r(t))^T$ and, for all $k \in \{1, \dots, r\}$, $N_k(t) \in \mathbb{Z}_{\geq 0}$ is the number of times the k th reaction has occurred up to time t . There are several established frameworks for modeling the time-evolution of CRNs on discrete state spaces, including that of continuous time Markov chains (CTMCs) and stochastic Petri nets. We will not concern ourselves with precise dynamical details; rather, we will focus on where trajectories may evolve in $\mathbb{Z}_{\geq 0}^m$. A similar treatment was conducted by Pauleve et al. (2014).

Note that the state space of a subconservative CRN is finite (Theorem 1 of Memmi and Roucairol 1975). For subconservative CRNs, therefore, the notion of recurrence introduced above agrees with the notion of positive recurrence from the language of CTMC (Lawler 2006).

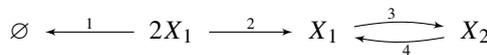
We now extend the properties of recurrence and transience of states to the complexes of a CRN.

Definition 3 Consider a CRN on a discrete state space. Then:

1. A complex $y \in \mathcal{C}$ is *strongly recurrent* from $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$ if, for any $\mathbf{Y} \in \mathbb{Z}_{\geq 0}^m$ such that $\mathbf{X} \rightsquigarrow \mathbf{Y}$, there is a $\mathbf{Z} \in \mathbb{Z}_{\geq 0}^m$ for which $\mathbf{Y} \rightsquigarrow \mathbf{Z}$ and y is charged at \mathbf{Z} ; otherwise, y is *weakly transient* from \mathbf{X} .
2. A complex $y \in \mathcal{C}$ is *weakly recurrent* from $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$ if there is a $\mathbf{Y} \in \mathbb{Z}_{\geq 0}^m$ such that $\mathbf{X} \rightsquigarrow \mathbf{Y}$ and y is strongly recurrent from \mathbf{Y} ; otherwise, y is *strongly transient* from \mathbf{X} .

In plain English, a complex y is strongly recurrent from a state if, no matter where the process goes, it can always reach a state where y is charged. A complex is y is weakly recurrent from a state if the process can reach a state from which y is strongly recurrent.

To show that a complex can be weakly recurrent (respectively, weakly transient) without being strongly recurrent (respectively, strongly transient), consider the following CRN:



For any state $\mathbf{X} = (X_1^\#, X_2^\#) \in \mathbb{Z}_{\geq 0}^2$ satisfying $X_1^\# + X_2^\# \geq 2$, we have that the complex $2X_1$ is strongly transient and the complexes X_1 and X_2 are both weakly recurrent and weakly transient but not strongly recurrent or strongly transient. This is due to the observation that there are sequences of reactions which can lead to either the state $\{X_1^\# = 0, X_2^\# = 0\}$, from which no complexes are ever charged, or to one of the states $\{X_1^\# = 1, X_2^\# = 0\}$ and $\{X_1^\# = 0, X_2^\# = 1\}$, from which both X_1 and X_2 are strongly recurrent.

The following clarifies the type of behavior for CRNs on discrete state spaces in which we will be interested.

Definition 4 Consider a CRN on a discrete state space. We will say that the CRN exhibits:

1. an *extinction event* on $\mathcal{Y} \subseteq \mathcal{C}$ from $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$ if every complex $y \in \mathcal{Y}$ is strongly transient from \mathbf{X} .
2. a *guaranteed extinction event* on $\mathcal{Y} \subseteq \mathcal{C}$ if it has an extinction event on \mathcal{Y} from every $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$.

Example 4 Consider the CRN (1) introduced in the introduction. Through repeated application of reaction 3, we can arrive at the state $\{X_1^\# = M, X_2^\# = 0\}$ where $M = X_1^\# + X_2^\#$. Since this is a possible outcome from any initial $\mathbf{X} \in \mathbb{Z}_{\geq 0}^2$, we have that this CRN has a guaranteed extinction event on $\mathcal{Y} = \{X_1 + X_2, 2X_2, X_2\}$. Notice that no reaction may occur after the extinction event.

Example 5 Consider the CRN in Example 3. Notice that the reaction $X_1 + X_2 \rightarrow X_1$ cannot occur indefinitely since all other reactions in the CRN preserve $X_1^\# + X_2^\#$. It follows that the model has a guaranteed extinction event on $\mathcal{Y} = \{X_1 + X_2\}$. Notice, however, that for every state $\mathbf{X} \in \mathbb{Z}_{\geq 0}^2$ satisfying $X_1^\# + X_2^\# \geq 1$ the complexes X_1 and X_2 are both strongly recurrent from \mathbf{X} . An extinction event therefore does not necessarily imply that all reactions must cease.

3 Main results

In this section, we motivate and present the main new constructions and theory of the paper (Lemma 2 and Theorem 1).

3.1 Domination-expanded reaction networks

We introduce the following.

Definition 5 Let $y, y' \in \mathcal{C}$ denote two distinct complexes of a CRN. We say that y *dominates* y' if $y' \leq y$. We define the *domination set* of a CRN to be

$$\mathcal{D}^* = \{(y, y') \in \mathcal{C} \times \mathcal{C} \mid y' \leq y, y \neq y'\}. \quad (6)$$

The notion of complex domination was introduced by Anderson et al. (2014) as an adaptation of the notion of “differing in one species” introduced by Shinar and Feinberg (2010). The domination property was extended to SLCs by Brijder (2015) who also showed that, for conservative CRNs, the domination properties give rise to a binary relation on the SLCs of a CRN whose transitive closure is a partial ordering on the SLCs of the CRN (Lemma 2 of Brijder 2015). We note that the definition of complex domination in Definition 5 is consistent with Brijder (2015) but reversed from Anderson et al. (2014).

Example 6 Consider the CRNs from Examples 1, 2, and 3 respectively. For the CRN in Example 1, we set $y_1 = X_1 + X_2, y_2 = 2X_2, y_3 = X_2$, and $y_4 = X_1$ and have $y_3 \leq y_1, y_3 \leq y_2$, and $y_4 \leq y_1$. For the CRN in Example 2, we set $y_1 = X_1, y_2 =$

$2X_2, y_3 = X_2$, and $y_4 = 2X_1$, and have $y_1 \leq y_4$ and $y_3 \leq y_2$. For the CRN in Example 3, we set $y_1 = X_1 + X_2, y_2 = X_1$, and $y_3 = X_2$, and have $y_2 \leq y_1$ and $y_3 \leq y_1$.

The key construction of this paper is the following, which uses the domination relations \leq to expand CRNs into larger CRNs we call *domination-expanded reaction networks*.

Definition 6 We say that $(\mathcal{S}, \mathcal{C}, \mathcal{R} \cup \mathcal{D})$ is a *domination-expanded reaction network* (dom-CRN) of the CRN $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ if $\mathcal{D} \subseteq \mathcal{D}^*$ and $\mathcal{R} \cap \mathcal{D} = \emptyset$. Furthermore, we say a dom-CRN is \mathcal{Y} -admissible if, given an absorbing complex set $\mathcal{Y} \subseteq \mathcal{C}$ of the dom-CRN, we have $(y, y') \in \mathcal{D}$ implies $y' \in \mathcal{Y}^c$.

A dom-CRN is a CRN which consists of the original CRN with additional directed edges corresponding to some (potentially all) of the domination relations $y' \leq y$. Note that the reaction arrows flow from the dominating complex to the “smaller” complex in the domination relation, i.e. $y' \leq y$ implies we add $y \rightarrow y'$. Like reactions, we will denote domination relations as either (y, y') or $y \rightarrow y'$. A dom-CRN is \mathcal{Y} -admissible if we do not add any reactions which lead to the absorbing complex set \mathcal{Y} of the dom-CRN.

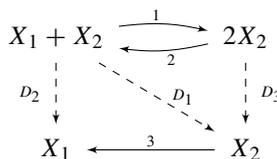
Remark 1 When applying Definition 6, we will commonly let the absorbing complex set \mathcal{Y} coincide with the set of terminal complexes of the dom-CRN, which we will show is a subset of the terminal complexes of original CRN in Lemma 1. In such cases, we will say a dom-CRN is simply *admissible* with the understanding that \mathcal{Y} is the set of terminal complexes.

Note that a dom-CRN is a CRN itself and therefore has associated to it all of the quantities and structural matrices given Sect. 2.1. While a dom-CRN in general may have different structural properties than the original CRN, an important restriction is given by the following result, which is based on Lemma 2 of Brijder (2015). The proof is contained in Appendix A.

Lemma 1 *If a CRN is subconservative, then for any dom-CRN: (i) the SLCs of the CRN and the dom-CRN coincide, and (ii) every terminal SLC of the dom-CRN is a terminal SLC of the CRN.*

We can interpret Lemma 1 as saying that, for a subconservative CRN, the addition of domination edges does not create new cycles between SLCs since this would create new SLCs.

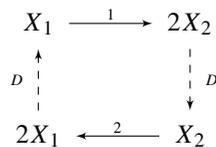
Example 7 Consider the CRN (1) from the Introduction and Examples 1 and 6. Recall that the CRN is conservative, and therefore subconservative, so that Lemma 1 applies. The dom-CRN with the maximal number of reactions is given by the following:



where we have indexed the domination relations for clarity. As guaranteed by Lemma 1, the SLCs of the CRN and dom-CRN coincide. Notice that the terminal complex X_1 in the dom-CRN above is terminal in the original CRN, but that the terminal complexes $X_1 + X_2$ and $2X_2$ in the CRN are not terminal in the dom-CRN.

Notice also that this dom-CRN is not admissible since the domination relations $X_1 + X_2 \rightarrow X_1$ leads to the terminal complex X_1 . Consider instead the subset $\mathcal{D} = \{X_1 + X_2 \rightarrow X_2, 2X_2 \rightarrow X_2\} \subset \mathcal{D}^*$. This generates the dom-CRN given as (2) in the Introduction. This dom-CRN is admissible since \mathcal{D} contains no domination edges which lead to the terminal complex X_1 .

Example 8 Consider the CRN from Examples 2 and 6. Recall that the CRN is neither conservative nor subconservative. Thus, Lemma 1 stands silent. The maximal dom-CRN is given by the following:



We have that there is only one SLC in the dom-CRN, which is given by $\{X_1, 2X_2, X_2, 2X_1\}$, so that the SLCs of the CRN and dom-CRN do not coincide. We can see, therefore, that the conclusion of Lemma 1 does not hold in general if we remove the subconservative assumption.

3.2 \mathcal{Y} -Exterior forests and balancing vectors

The following concept is adapted from numerous sources in graph theory. Directed rooted trees have been used extensively in CRNT (Craciun et al. 2009; Johnston 2014) and the related notion of arborescences was utilized in Boros (2013). Note that a forest is a graph which can be formed as the union of disjoint trees.

Definition 7 Consider a CRN $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ and a \mathcal{Y} -admissible dom-CRN $(\mathcal{S}, \mathcal{C}, \mathcal{R} \cup \mathcal{D})$ where $\mathcal{Y} \subseteq \mathcal{C}$ is an absorbing complex set on the dom-CRN. Then $(\mathcal{S}, \mathcal{C}, \mathcal{R}_F \cup \mathcal{D}_F)$ where $\mathcal{R}_F \subseteq \mathcal{R}$ and $\mathcal{D}_F \subseteq \mathcal{D}$ is called an \mathcal{Y} -exterior forest if, for every complex $y \in \mathcal{Y}^c$, there is a unique path in $\mathcal{R}_F \cup \mathcal{D}_F$ from y to \mathcal{Y} .

A \mathcal{Y} -exterior forest is a forest in the usual sense in graph theory after contracting \mathcal{Y} to a single point in the dom-CRN. Note that Definition 7 places no restrictions on \mathcal{Y} -interior reactions. By convention, we will include all such reactions in every \mathcal{Y} -exterior forest. If \mathcal{Y} consists solely of the terminal complexes of the dom-CRN, we say $(\mathcal{S}, \mathcal{C}, \mathcal{R}_F \cup \mathcal{D}_F)$ is simply an exterior forest.

We will be interested in particular in \mathcal{Y} -exterior forests which satisfy the following property.

Definition 8 Consider a CRN $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ and a \mathcal{Y} -admissible dom-CRN $(\mathcal{S}, \mathcal{C}, \mathcal{R} \cup \mathcal{D})$ where $\mathcal{Y} \subseteq \mathcal{C}$ is an absorbing complex set on the dom-CRN. Let \mathcal{D} be ordered so that

$\mathcal{D} = (D_1, \dots, D_d)$ where $d = |\mathcal{D}|$, and let Γ be the stoichiometric matrix associated with the original CRN. Then a \mathcal{Y} -exterior forest $(\mathcal{S}, \mathcal{C}, \mathcal{R}_F \cup \mathcal{D}_F)$ is said to be *balanced* if there is a vector $\alpha = (\alpha_R, \alpha_D) \in \mathbb{Z}_{\geq 0}^{r+d}$ with $\alpha_k > 0$ for at least one \mathcal{Y} -exterior reaction which satisfies:

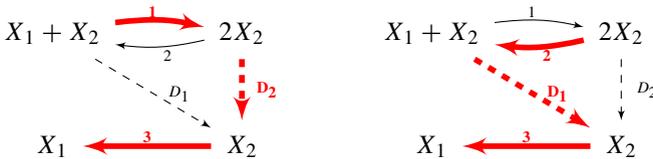
1. $\text{supp}(\alpha_R) \subseteq \text{supp}(\mathcal{R}_F)$ and $\text{supp}(\alpha_D) \subseteq \text{supp}(\mathcal{D}_F)$;
2. $\alpha_R \in \ker(\Gamma)$; and
3. for every $R_k = y \rightarrow y' \in \mathcal{R}_F \cup \mathcal{D}_F$ where $y \in \mathcal{Y}^c$, we have $\alpha_k \geq \sum_{R_l \in \Theta(y)} \alpha_l$ where $\Theta(y) = \{R_l \in \mathcal{R}_F \cup \mathcal{D}_F \mid R_l = y'' \rightarrow y\}$.

Otherwise, the \mathcal{Y} -exterior forest is said to be *unbalanced*.

The third condition of Definition 8 can be interpreted as saying that, for every $y \in \mathcal{Y}^c$, the weight of the outgoing edge in the \mathcal{Y} -exterior forest must be at least as large as the sum of all incoming edges.

When taken together, the three conditions of Definition 8 generate a set of equalities and inequalities on the edges of the dom-CRN. This suggests a computational implementation. Such an implementation is presented in the companion paper by Johnston (2017) where the conditions of Definitions 7 and 8 are checked on 458 models from the European Bioinformatics Institute’s BioModels Database. The program is implemented using a series of linear programming modules.

Example 9 Recall the CRN (1) taken from the introduction, Examples 1, 6, and 7, and the admissible dom-CRN from Example 7. This dom-CRN admits several exterior forests, for example the following substructures in bold red:



Note that every nonterminal complex has a unique path to X_1 . We now check whether these exterior forests are balanced by Definition 8 by checking equalities and inequalities on the vector of edges of the following form:

$$\begin{array}{c} \text{reaction: } 1 \quad 2 \quad 3 \quad D_1 \quad D_2 \\ \hline \alpha = ((\alpha_R)_1, (\alpha_R)_2, (\alpha_R)_3, (\alpha_D)_1, (\alpha_D)_2) \end{array}$$

$\underbrace{\hspace{10em}}_{\alpha_R} \qquad \underbrace{\hspace{5em}}_{\alpha_D}$

1. In order for the left exterior forest to be balanced, it is required that we find a vector $\alpha = ((\alpha_R)_1, (\alpha_R)_2, (\alpha_R)_3, (\alpha_D)_1, (\alpha_D)_2) \in \mathbb{R}_{\geq 0}^5$, $\alpha \neq \mathbf{0}$, satisfying:

$$\left\{ \begin{array}{l} \text{(Cond. 1):} \quad (\alpha_R)_2 = 0, (\alpha_D)_1 = 0 \\ \text{(Cond. 2):} \quad -(\alpha_R)_1 + (\alpha_R)_2 + (\alpha_R)_3 = 0 \\ \qquad \qquad \qquad (\alpha_R)_1 - (\alpha_R)_2 - (\alpha_R)_3 = 0 \\ \text{(Cond. 3):} \quad (\alpha_R)_3 \geq (\alpha_D)_2 \geq (\alpha_R)_1 \geq 0. \end{array} \right.$$

We can choose $(1, 0, 1, 0, 1)$ so that this is balanced exterior forest.

- In order for the right exterior forest to be balanced, it is required that we find a nontrivial vector $\alpha = ((\alpha_R)_1, (\alpha_R)_2, (\alpha_R)_3, (\alpha_D)_1, (\alpha_D)_2) \in \mathbb{R}_{\geq 0}^5, \alpha \neq \mathbf{0}$, satisfying:

$$\left\{ \begin{array}{l} \text{(Cond. 1):} \quad (\alpha_R)_1 = 0, (\alpha_D)_2 = 0 \\ \text{(Cond. 2):} \quad -(\alpha_R)_1 + (\alpha_R)_2 + (\alpha_R)_3 = 0 \\ \quad \quad \quad (\alpha_R)_1 - (\alpha_R)_2 - (\alpha_R)_3 = 0 \\ \text{(Cond. 3):} \quad (\alpha_R)_3 \geq (\alpha_D)_1 \geq (\alpha_R)_2 \geq 0. \end{array} \right.$$

Substituting Condition 1 into Condition 2 gives $(\alpha_R)_2 + (\alpha_R)_3 = 0$ which is inconsistent with the requirement from Condition 3 that $(\alpha_R)_3 \geq (\alpha_R)_2 \geq 0$ and at least one entry be nonzero. It follows that this is an unbalanced exterior forest.

3.3 Conditions for extinction events

We now present the main results of this paper, which are inspired by Theorem 1 and Corollary 2 of [Brijder \(2015\)](#). The proof of Lemma 2 is contained in Appendix B.

Lemma 2 *Consider a subconservative CRN and a \mathcal{Y} -admissible dom-CRN where $\mathcal{Y} \subseteq \mathcal{C}$ is an absorbing complex set on the dom-CRN. Suppose that there is a complex $y \in \mathcal{Y}^c$ of the dom-CRN which is weakly recurrent from a state $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$ in the discrete state space CRN. Then every \mathcal{Y} -exterior forest of the dom-CRN is balanced.*

This result places restrictions on the structure of a subconservative CRN that does not experience a guaranteed extinction event. We will be more frequently interested in when discrete extinction occurs, and therefore present the following result which follows immediately as the contrapositive of Lemma 2.

Theorem 1 *Consider a subconservative CRN and a \mathcal{Y} -admissible dom-CRN where $\mathcal{Y} \subseteq \mathcal{C}$ is an absorbing complex set on the dom-CRN. Suppose there is a \mathcal{Y} -exterior forest of the dom-CRN which is unbalanced. Then the discrete state space CRN has a guaranteed extinction event on \mathcal{Y}^c .*

Recall that an exterior forest is unbalanced if there is a set of equalities and inequalities on the edges of the dom-CRN which cannot be satisfied. The question of determining sufficient conditions for discrete extinction is therefore reduced to determining the feasibility of particular sets of equalities and inequalities.

Notice also that, even if a CRN permits many \mathcal{Y} -exterior forests, it is sufficient for a single one to be unbalanced for an extinction event to follow. Furthermore, the set of strongly transient complexes corresponds to the set of complexes not in \mathcal{Y} . Note that this may contain terminal complexes in the original CRN (see Example 7).

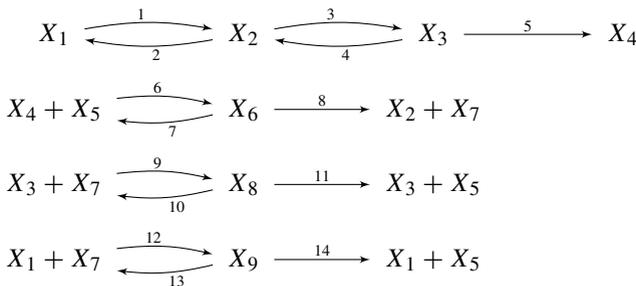
Remark 2 By convention, when applying Theorem 1, if no mention of an absorbing complex set $\mathcal{Y} \subseteq \mathcal{C}$ is made, it is assumed to be the set of terminal complexes in the dom-CRN.

Example 10 Reconsider the CRN analyzed in Examples 1, 6, and 7. This CRN is conservative, and in Example 9 we showed that there is an admissible dom-CRN with an unbalanced exterior forest. It follows from Theorem 1 that the discrete state space CRN has a guaranteed extinction event on the set of nonterminal complexes of the dom-CRN. That is, from all states $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$, there is guaranteed to be a time after which the count of the species is insufficient for any reaction from the complexes $X_1 + X_2$, $2X_2$, and X_2 to occur. This is consistent with our earlier observation that the state $\{X_1^\# = M, X_2^\# = 0\}$ where $M = X_1^\# + X_2^\#$ absorbs all trajectories through repeated application of the reactions $2X_2 \rightarrow X_1 + X_2$ and $X_2 \rightarrow X_1$. Notice that this pathway consists of the true reactions in the unbalanced exterior forest.

3.4 EnvZ-OmpR signaling pathway

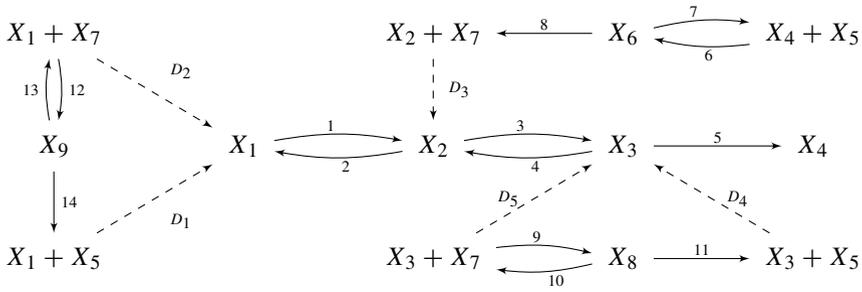
In this section, we consider a CRN which was proposed as underlying the EnvZ/OmpR signaling pathway in *Escherichia coli* in Shinar and Feinberg (2010). This CRN has been studied previously with a discrete state space by Anderson et al. (2014) and Brijder (2015) where it was shown to exhibit a guaranteed extinction event.

Example 11 Consider the following reaction mechanism, which was proposed by Shinar and Feinberg (2010) as underlying the EnvZ/OmpR signaling pathway in *E. coli* in the Supplemental Material:

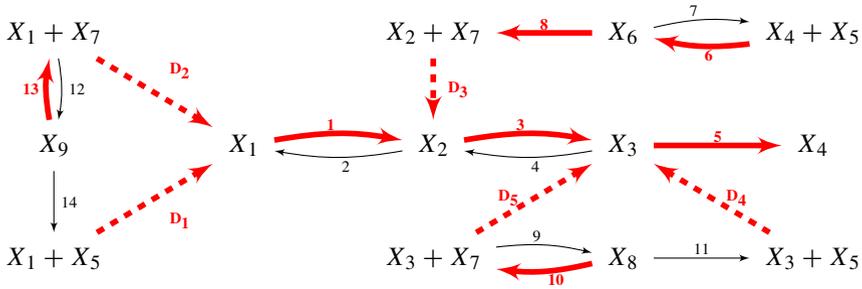


where $X_1 = EnvZ-ADP$, $X_2 = EnvZ$, $X_3 = EnvZ-ATP$, $X_4 = EnvZ_p$, $X_5 = OmpR$, $X_6 = EnvZ_p-OmpR$, $X_7 = OmpR_p$, $X_8 = EnvZ-ATP-OmpR_p$, $X_9 = EnvZ-ADP-OmpR_p$.

Consider the admissible dom-CRN with $\mathcal{D} = \{X_1 + X_5 \xrightarrow{D_1} X_1, X_1 + X_7 \xrightarrow{D_2} X_1, X_2 + X_7 \xrightarrow{D_3} X_2, X_3 + X_5 \xrightarrow{D_4} X_3, X_3 + X_7 \xrightarrow{D_5} X_3\}$. The dom-CRN may be graphically represented as:



Consider furthermore the following exterior forest:



In the highlighted structure (bold red), there is a unique path from every complex to the terminal complex X_4 . It can be seen directly that this exterior forest is unbalanced by noting that we need a vector $\alpha = (\alpha_R, \alpha_D) \in \mathbb{Z}_{\geq 0}^{19}, \alpha \neq \mathbf{0}$, which has support on a subset of the red highlighted structure above. To satisfy Condition 2 of Definition 8, we need to satisfy $\alpha_R \in \ker(\Gamma)$. We can check that $\ker(\Gamma) \cap \mathbb{R}_{\geq 0}^r$ has the generators:

$$\begin{array}{c}
 \text{reaction: } 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10 \ 11 \ 12 \ 13 \ 14 \\
 \hline
 \ker(\Gamma) \cap \mathbb{R}_{\geq 0}^r = \{ (1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), \\
 (0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), \\
 (0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0), \\
 (0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0), \\
 (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0), \\
 (0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0), \\
 (0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1) \}
 \end{array}$$

The first five vectors correspond to reversible reaction pairs in the CRN and so may be ignored. In order to obtain a nontrivial vector α_R , we require $(\alpha_R)_5 > 0$. To build such a vector using the sixth vector yields a vector with support on $(\alpha_R)_{11}$ while building it out of the seventh vector yields a vector with support on $(\alpha_R)_{14}$. Neither of these options is consistent with Condition 1 of Definition 8 so that the exterior forest is unbalanced. It follows by Theorem 1 that the discrete state space CRN has a guaranteed extinction event, and that every complex except X_4 is strongly transient.

In fact, all trajectories are absorbed by a state where $X_4^\# > 0$, $X_7^\# > 0$, and $X_i^\# = 0$ for $i \in \{1, 2, 3, 5, 6, 8, 9\}$.

This result was previously obtained in the Supplemental Online Material of [Anderson et al. \(2014\)](#) and also proved for a simplified CRN by [Brijder \(2015\)](#). Notably, our method of constructing a dom-CRN suggests pathways toward extinction by restricting to the reactions in the unbalanced exterior forest. We have the following pathway:

1. Fire reactions 10 and 13 until $X_8^\# = 0$ and $X_9^\# = 0$.
2. Fire reaction 1 until $X_1^\# = 0$.
3. Fire reactions 6 and 8 until $X_6^\# = 0$ and either $X_4^\# = 0$ or $X_5^\# = 0$.
4. Fire reactions 3 and 5 until $X_2^\# = 0$ and $X_3^\# = 0$.
5. Repeat steps 3 and 4 until $X_5^\# = 0$.

Notice that the sequences of reactions in steps 3 and 4 convert X_2 into X_4 and vice versa, but that X_5 is irreversibly converted into X_7 . It follows that $X_5^\# = 0$ will eventually be attained and consequently that the algorithm will terminate.

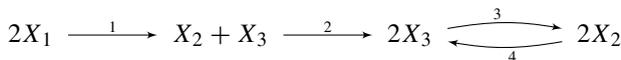
A similar pathway was constructed by [Anderson et al. \(2014\)](#) for a smaller model but was not apparent by the main result (Theorem 4) itself. We have constructed the pathway here by firing the extremal reactions in the \mathcal{Y} -exterior forest to exhaustion, and tracking which species disappears.

3.5 Importance of technical conditions

In this section, we provide further examples which demonstrate how to apply Theorem 1, and which demonstrate the necessity of several of the technical conditions required of the result.

Example 12 presents a CRN which can be shown to have an extinction event for an absorbing complex set $\mathcal{Y} \subseteq \mathcal{C}$ which is not the set of terminal complexes in the dom-CRN. Example 13 presents a CRN which does not have a guaranteed extinction event, but which can be shown to have an unbalanced exterior forest if we do not insist on the underlying dom-CRN being admissible. Example 14 demonstrates that including Condition 3 of Definition 8 allows further classification of CRNs with extinction events than would be possible otherwise.

Example 12 It is natural to wonder whether, when applying Theorem 1, there is an advantage to generalizing the set of terminal complexes to an absorbing complex set $\mathcal{Y} \subseteq \mathcal{C}$. To show that there is, consider the following CRN:



There are no domination relations so that the only dom-CRN corresponds to the CRN shown, and it is trivially admissible. The only exterior forest consists of all reactions. Notable, it contains reactions 1 and 2 on the nonterminal component. We can easily determine that $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (0, 2, 1, 0)$ satisfies the conditions of Definition 8 and therefore that this exterior forest is balanced. Therefore, Theorem 1 does not apply and we may not conclude that an extinction event occurs.

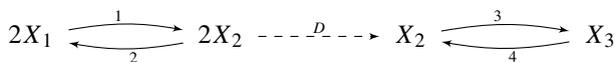
Consider instead taking $\mathcal{Y} = \{X_2 + X_3, 2X_3, 2X_2\}$. This set is absorbing and contains every terminal complex of the CRN. The only exterior forest again contains all reactions but only reaction 1 is \mathcal{Y} -exterior. Since there is no balancing vector α for which $\alpha_1 \neq 0$, we may conclude by Theorem 1 that there is a guaranteed extinction event on $\mathcal{Y}^c = \{2X_1\}$. In fact, we can see this directly since repeated application of reaction 1 will deplete X_1 and there are no pathways by which to replenish it.

Example 13 It is natural to wonder whether it is necessary to insist on dom-CRNs being \mathcal{Y} -admissible. To show that removing this assumption from Theorem 1 can lead to misclassification, consider the following CRN:



The CRN has only the single domination relation $X_2 \leq 2X_2$. Since the corresponding domination relation $2X_2 \rightarrow X_2$ leads to a terminal component in any resulting dom-CRN, we may not add it, so that the only admissible dom-CRN corresponds to the original CRN.

Suppose, however, that we do not insist on dom-CRNs being admissible. Specifically, suppose we allow the following dom-CRN:



The only exterior forest is given in bold red as follows:

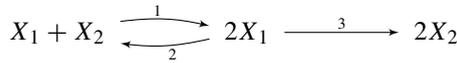


Notice that we have included the terminal reactions in the exterior forest. In order to be balanced, we must find a vector $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_D)$ which is nonzero on at least one of the nonterminal reactions α_1 and α_2 , such that

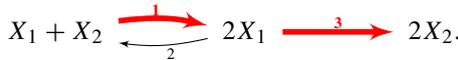
$$\left\{ \begin{array}{l} \text{(Cond. 1) : } \alpha_2 = 0 \\ \text{(Cond. 2) : } -2\alpha_1 + 2\alpha_2 = 0 \\ \qquad \qquad \qquad 2\alpha_1 - 2\alpha_2 - \alpha_3 + \alpha_4 = 0 \\ \qquad \qquad \qquad \alpha_3 - \alpha_4 = 0 \\ \text{(Cond. 3) : } \alpha_D \geq \alpha_1 \geq 0. \end{array} \right.$$

Conditions 1 and 2 imply that $\alpha_1 = 0$ so that α does not have support on the non-terminal portion of the dom-CRN. It follows that the exterior forest is unbalanced. Note, however, that Theorem 1 remains silent since the presented dom-CRN is not admissible. There is, however, clearly no extinction event in this CRN since all reactions of the discrete state space CRN are strongly recurrence from any state $\mathbf{X} \in \mathbb{Z}_{\geq 0}^3$ satisfying $X_1^\# + X_2^\# + X_3^\# \geq 3$. This example therefore highlights the importance of the assumption that dom-CRNs be \mathcal{Y} -admissible.

Example 14 It is natural to wonder whether Condition 3 of Definition 8 is useful in classifying discrete state space CRNs with extinction events. To see that it can be, consider the following CRN:



There are no domination relations so the dom-CRN coincides with the original CRN. We have only the following exterior forest in bold red:



In order for this exterior forest to be balanced, we need to have a vector $\alpha = (\alpha_1, \alpha_2, \alpha_3), \alpha \neq \mathbf{0}$, which satisfies the following equalities and inequalities:

$$\left\{ \begin{array}{l} \text{(Cond. 1) : } \alpha_2 = 0 \\ \text{(Cond. 2) : } \alpha_1 - \alpha_2 - 2\alpha_3 = 0 \\ \qquad \qquad -\alpha_1 + \alpha_2 + 2\alpha_3 = 0 \\ \text{(Cond. 3) : } \alpha_3 \geq \alpha_1 \geq 0. \end{array} \right.$$

Condition 1 reduces Condition 2 to $\alpha_1 = 2\alpha_3$, so that, combining with Condition 3, we have

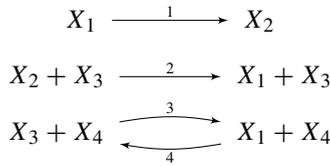
$$\alpha_3 \geq \alpha_1 = 2\alpha_3 \geq 0.$$

This can only be satisfied by $\alpha_1 = 0$ and $\alpha_3 = 0$, which is a violation of the requirement that α be nonzero. It follows that the exterior forest is unbalanced and therefore, by Theorem 1, the discrete state space CRN has a guaranteed extinction event on the nonterminal complexes $X_1 + X_2$ and $2X_1$. Note, however, that the vector $\alpha = (2, 0, 1)$ satisfies Conditions 1 and 2 of Definition 8. It follows that Condition 3 of Definition 8 allows furthermore classification of CRNs with extinction events than Conditions 1 and 2 allow by themselves. Note also that this CRN is not classified as having a guaranteed extinction event by Corollary 2 of Brijder (2015).

3.6 Conditions are sufficient but not necessary

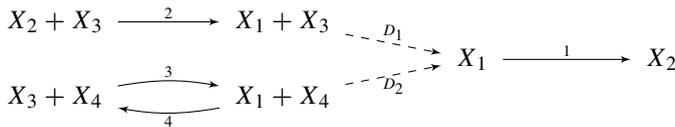
It is natural to wonder whether the conditions of Lemma 2 and Theorem 1 are necessary as well as sufficient for a discrete state space CRN to have an extinction event. Examples 15 and 16 show that the conditions are sufficient only.

Example 15 Consider the following CRN:

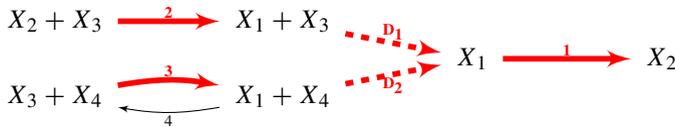


The CRN has a guaranteed discrete extinction event, since X_3 may convert into X_1 through reaction 3, then X_1 may convert into X_2 through reaction 1. This shuts down all reactions.

To show that Theorem 1 is incapable of affirming this extinction event, it is necessary to show that every \mathcal{Y} -exterior forest of every \mathcal{Y} -admissible dom-CRN is balanced. We start by considering the terminal complexes and the set $\mathcal{D} = \{X_1 + X_3 \xrightarrow{D_1} X_1, X_1 + X_4 \xrightarrow{D_2} X_1\}$. This gives the following dom-CRN:



This dom-CRN is admissible and admits only a single exterior forest in bold red:



This forest is balanced if we have a nontrivial vector $\alpha = ((\alpha_R)_1, (\alpha_R)_2, (\alpha_R)_3, (\alpha_R)_4, (\alpha_D)_1, (\alpha_D)_2)$, $\alpha_R \neq \mathbf{0}$, which satisfies the following:

$$\left\{ \begin{array}{l}
 \text{(Cond. 1):} \quad (\alpha_R)_4 = 0 \\
 \text{(Cond. 2):} \quad -(\alpha_R)_1 + (\alpha_R)_2 + (\alpha_R)_3 - (\alpha_R)_4 = 0 \\
 \quad \quad \quad (\alpha_R)_1 - (\alpha_R)_2 = 0 \\
 \quad \quad \quad -(\alpha_R)_3 + (\alpha_R)_4 = 0 \\
 \text{(Cond. 3):} \quad (\alpha_D)_1 + (\alpha_D)_2 \geq (\alpha_R)_1 \geq 0 \\
 \quad \quad \quad (\alpha_D)_1 \geq (\alpha_R)_2 \\
 \quad \quad \quad (\alpha_D)_2 \geq (\alpha_R)_3.
 \end{array} \right. \tag{7}$$

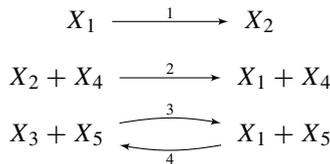
This can be satisfied by the vector $\alpha = (1, 1, 0, 0, 1, 0)$. It follows that the forest is balanced, and since this is the only exterior forest for the given dom-CRN, no conclusion may be reached as a result of Theorem 1.

We now consider more general absorbing complex sets $\mathcal{Y} \subseteq \mathcal{C}$. Notice that any potential \mathcal{Y} which contains a subset of $\{X_2, X_1, X_2 + X_3, X_1 + X_3\}$ can be balanced by the α above, with perhaps different support on α_D . If $X_1 \in \mathcal{Y}$, however, we must have $\mathcal{D} = \emptyset$ in order for the dom-CRN to be \mathcal{Y} -admissible. Otherwise, we would have

an edge in \mathcal{D} which would lead to \mathcal{Y} . For $\mathcal{D} = \emptyset$, however, we have that $X_3 + X_4$ and $X_1 + X_4$ are terminal in the dom-CRN and therefore $X_3 + X_4$ and $X_1 + X_4$ must be included in \mathcal{Y} . This leaves $\mathcal{Y} = \mathcal{C}$ which has an empty exterior forest. There are no other cases to consider, so we are done.

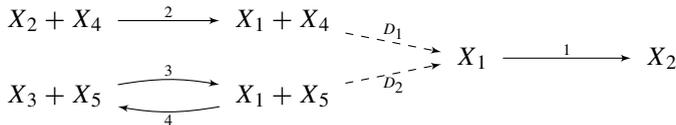
It follows that every \mathcal{Y} -exterior forest of every \mathcal{Y} -admissible dom-CRN is balanced. Since the CRN has a guaranteed extinction event, however, it follows that the conditions of Theorem 1 are not necessary for extinction events in discrete state space CRNs.

Example 16 To show that the gap raised in Example 15 may not be easily overcome by structural considerations alone, consider the following CRN:



This is the CRN in Example 15 with X_3 replaced with X_4 in the reaction 2, and X_4 replaced with X_5 in reactions 3 and 4. Examples 15 and 16 share significant structural data, including connectivity of paths, domination relations between complexes, and $\ker(\Gamma)$.

Taking $\mathcal{D} = \{X_1 + X_4 \rightarrow X_1, X_1 + X_5 \rightarrow X_1\}$ gives the following admissible dom-CRN:



We arrive at the same balancing equalities and inequalities (7) as Example 15, so that every \mathcal{Y} -exterior forest on this dom-CRN is balanced. Since the connectivity and domination relations are shared with Example 15, we can exhaust nontrivial \mathcal{Y} -admissible dom-CRNs in the same way, and we conclude that Theorem 1 is inconclusive.

In contrast to Example 15, this example does not exhibit an extinction event for most initial conditions. Every complex is strongly recurrent from every state $\mathbf{X}_{\geq 0}^5$ such that $X_4^\# > 0$, $X_5^\# > 0$, and any one of $X_1^\#, X_2^\#,$ and $X_3^\#$ is positive. This analysis suggests that comprehensive conditions for extinction events must depend on further structural information than that considered in this paper.

4 Conclusions and future work

In this paper, we have presented novel conditions (Lemma 2 and Theorem 1) on the structure of a CRN that are sufficient to guarantee that the corresponding CRN exhibits an extinction event. The conditions presented generalize the dependence on terminal SLCs in Anderson et al. (2014) and Brijder (2015), and also produces a

system of equalities and inequalities which can be directly verified. Our conditions are fundamentally graphic-theoretical in nature and suggest pathways to extinction.

The results of this paper have consequences for the well-studied area of CTMC models of biochemical reaction networks. In particular, the extinctions guaranteed to exist by the analysis presented here are often rare events on relevant timescales. In such situations it is their quasi-stationary distributions, as opposed to their stationary distributions, that must be characterized to gain insight into model behavior. Analyses on the nature of these distributions has been conducted by [Anderson et al. \(2014, 2017\)](#), [Enciso \(2016\)](#).

This work raises several promising avenues for future work:

1. While [Theorem 1](#) gives sufficient conditions for discrete extinction, they are not necessary (see [Examples 15 and 16](#)). This raises the question of whether there are structural conditions which are both sufficient and necessary for discrete extinction and, if so, which further structural components of the CRN might be utilized in such a result.
2. The conditions of [Theorem 1](#) consist of a system of equalities and inequalities. This suggests a computational implementation amenable, in particular, to the methods of linear programming. Linear programming has already been used widely in CRNT for verifying CRNs with desirable structural properties when desirable structural properties are present ([Johnston et al. 2012, 2016](#); [Johnston 2014, 2016](#); [Szederkényi 2010](#)). This computational framework is explored and utilized to characterize CRNs with extinction events in the companion paper by [Johnston \(2017\)](#).

Acknowledgements MDJ and DFA were supported by Army Research Office Grant W911NF-14-1-0401. DFA was also supported by NSF-DMS-1318832 and MDJ was also supported by the Henry Woodward Fund. GC was supported by NSF-DMS-1412643. RB is a postdoctoral fellow of the Research Foundation—Flanders (FWO). The authors are also grateful to the anonymous referees whose suggestions have greatly improved the readability, clarity, and accuracy of the paper.

Appendix A: Proof of Lemma 1

Proof of (i) Consider a subconservative CRN and dom-CRN. Since the reactions of the CRN are contained in the reactions of the dom-CRN, it follows that the SLCs of CRN remain strongly connected in the dom-CRN and therefore are contained in the SLCs of the dom-CRN.

Now suppose that there is an SLC of the dom-CRN which is not contained in any SLC of the CRN. It follows that there are SLCs $W, W' \in \mathcal{W}$ of the CRN such that is a path in the dom-CRN from some complex $y_0 \in W$ to some complex $y'_0 \in W'$, and there is a path in the dom-CRN from some complex $y'_1 \in W'$ to some complex $y_1 \in W$. Since W and W' are strongly connected, we can create a cycle in the dom-CRN by constructing a path from y_0 to y'_0 to y'_1 to y_1 back to y_0 . Furthermore, since this is not a cycle in the CRN (otherwise, W and W' would not be maximally strongly connected in the CRN), we have that there is at least one reaction in this cycle which is from \mathcal{D} .

We now index the complexes in the cycle so that, if there are $d' \geq 1$ reactions from \mathcal{D} in the cycle, we have the following segments in between these reactions:

$$\begin{aligned}
 & y_1^{(1)} \rightarrow y_2^{(1)} \rightarrow \dots \rightarrow y_{n_1}^{(1)} \\
 & y_1^{(2)} \rightarrow y_2^{(2)} \rightarrow \dots \rightarrow y_{n_2}^{(2)} \\
 & \quad \quad \quad \vdots \\
 & y_1^{(d')} \rightarrow y_2^{(d')} \rightarrow \dots \rightarrow y_{n_{d'}}^{(d')}.
 \end{aligned}
 \tag{8}$$

By construction, the segments above are connected by reactions in \mathcal{R} and satisfy $y_1^{(i+1)} \leq y_{n_i}^{(i)}$ and $y_1^{(1)} \leq y_{n_{d'}}^{(d')}$.

Let $\alpha \in \mathbb{Z}_{\geq 0}^r$ denote the vector of counts of the reactions in (8), and define $y_1^{(d'+1)} = y_1^{(1)}$. It follows that

$$\begin{aligned}
 \Gamma\alpha &= \sum_{i=1}^{d'} \sum_{j=1}^{n_i-1} (y_{j+1}^{(i)} - y_j^{(i)}) \\
 &= \sum_{i=1}^{d'} (y_{n_i}^{(i)} - y_1^{(i)}) \\
 &= \sum_{i=1}^{d'} (y_{n_i}^{(i)} - y_1^{(i+1)}) \geq \mathbf{0}
 \end{aligned}
 \tag{9}$$

by the domination relations $y_1^{(i+1)} \leq y_{n_i}^{(i)}$. Since the CRN is subconservative, it follows that there is a $\mathbf{c} \in \mathbb{R}_{>0}^m$ such that $\mathbf{c}^T \Gamma \leq \mathbf{0}$. It follows that we have

$$0 \leq [\mathbf{c}^T \Gamma]\alpha = \mathbf{c}^T [\Gamma\alpha] > 0$$

where the last strict inequality follows from $c_i > 0$ for $i \in \{1, \dots, m\}$ and the observation that at least one component in (9) must be strictly greater than zero since the complexes of the CRN are stoichiometrically distinct. This is a contradiction. It follows that such a cycle does not exist in the dom-CRN so that W and W' are SLCs of the CRN. The SLCs of the CRN and dom-CRN therefore coincide and (i) is shown. □

Proof of (ii) Note that (i) guarantees that the CRN and dom-CRN share the same set of SLCs which we will denote \mathcal{W} . Suppose that $W \in \mathcal{W}$ is terminal in the dom-CRN but not in the CRN. This implies that there is a reaction $(y, y') \in \mathcal{R}$ where $y \in W$ and $y' \notin W$; however, this reaction is included in the dom-CRN so that W may not be terminal in the dom-CRN. It follows that every terminal SLC of the dom-CRN is a terminal SLC of the CRN, and (ii) is shown. □

Appendix B: Proofs of Lemma 2 and Theorem 1

Remark 3 The following proof is inspired by the proof of Theorem 1 of [Brijder \(2015\)](#). The notation has been adapted to that of CRNT.

Proof Consider a subconservative CRN and a \mathcal{Y} -admissible dom-CRN where $\mathcal{Y} \subseteq \mathcal{C}$ is an absorbing complex set on the dom-CRN. Suppose that there is a complex $y \in \mathcal{Y}^c$ of the dom-CRN which is weakly recurrent from a state $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$. We will show that every \mathcal{Y} -exterior forest is balanced; that is, every \mathcal{Y} -exterior forest admits a vector $\alpha = (\alpha_R, \alpha_D) \in \mathbb{R}_{\geq 0}^{r+d}$ satisfying the requirements of Definition 8. We will accomplish this by constructing a sequence of reactions which may be executed indefinitely, and then demonstrating that this sequence repeats. We will define α based on a specific repeating portion of this sequence and show that it is balanced.

Let $\mathbf{X} \in \mathbb{Z}_{\geq 0}^m$ denote our initial state. By the assumption of weak recurrence, there is a sequence of reactions from \mathbf{X} to \mathbf{X}^0 such that there is a $y \in \mathcal{Y}^c$ which is strongly recurrent from \mathbf{X}^0 . It follows from this strong recurrence that there is a state $\mathbf{X}^{1-} \in \mathbb{Z}_{\geq 0}^m$ and a complex $y^{1-} \in \mathcal{Y}^c$ such that (i) $\mathbf{X}^0 \rightsquigarrow \mathbf{X}^{1-}$, (ii) y^{1-} is charged at \mathbf{X}^{1-} , and (iii) no complex $y \in \mathcal{Y}^c$ is charged at any state along the sequence of reactions from $\mathbf{X} \rightsquigarrow \mathbf{X}^{1-}$ except \mathbf{X}^{1-} . That is, y^{1-} is the first \mathcal{Y} -exterior complex which becomes charged as a result of the reaction sequence, and it is first charged at the state \mathbf{X}^{1-} . Note that, the first two conditions follow immediately from the recurrence assumption, and the third follows by taking any sequence guaranteed by strong recurrence, truncating at the first \mathcal{Y} -exterior complex which becomes charged, and redefining $y^{1-} \in \mathcal{Y}^c$ accordingly. Note also that, if y^{1-} is charged at \mathbf{X} originally, then the sequence of reactions is empty.

By construction, there is a unique path in the exterior forest from y^{1-} to \mathcal{Y} . Let $y^{1+} \in \mathcal{Y}$ denote the complex at the end of this path and $\mathbf{X}^{1+} \in \mathbb{Z}_{\geq 0}^m$ denote the state obtained by the sequential occurrence of the true reactions in the path (i.e. include reactions in \mathcal{R}_F but exclude domination relations \mathcal{D}_F). Note that (i) $\mathbf{X}^{1-} \rightsquigarrow \mathbf{X}^{1+}$, (ii) y^{1+} is charged at \mathbf{X}^{1+} , and (iii) this path contains at least one reaction in \mathcal{R}_F (i.e. the sequence is nonempty). The third property follows from the observation that $y^{1-} \in \mathcal{Y}^c$ and $y^{1+} \in \mathcal{Y}$, and the assumption that the \mathcal{Y} -exterior forest is admissible and therefore the last reaction in any \mathcal{Y} -exterior path to \mathcal{Y} is in \mathcal{R}_F .

We now iterate this procedure for $i = 2, 3, 4, \dots$, starting from the state $\mathbf{X}^{(i-1)+}$ rather than \mathbf{X} . This generates the following sequence of transitions, which may be continued indefinitely because there is a complex $y \in \mathcal{Y}^c$ which is strongly recurrent from \mathbf{X}^0 and the construction of the \mathcal{Y} -exterior forest:

$$\mathbf{X}^0 \rightsquigarrow \mathbf{X}^{1-} \rightsquigarrow \mathbf{X}^{1+} \rightsquigarrow \mathbf{X}^{2-} \rightsquigarrow \mathbf{X}^{2+} \rightsquigarrow \dots \tag{10}$$

Since the CRN is subconservative, we have that there is a finite number of accessible states (Theorem 1 of [Memmi and Roucairol 1975](#)). It follows that there is a state in $\{\mathbf{X}^{1-}, \mathbf{X}^{2-}, \dots\}$ which is repeated. We let n_1 and n_2 where $0 < n_1 < n_2$ denote the first and second indices for the set $\{\mathbf{X}^{1-}, \mathbf{X}^{2-}, \dots\}$ such that $\mathbf{X}^{n_1-} = \mathbf{X}^{n_2-}$. This gives the following subsequence of (10)

$$\mathbf{X}^{n_1-} \rightsquigarrow \mathbf{X}^{n_1+} \rightsquigarrow \dots \rightsquigarrow \mathbf{X}^{(n_2-1)+} \rightsquigarrow \mathbf{X}^{n_2-}. \tag{11}$$

Since $\mathbf{X}^{n_1-} = \mathbf{X}^{n_2-}$, (11) defines a sequence of reactions which can be repeated indefinitely.

We now define the vector $\alpha = (\alpha_R, \alpha_D) \in \mathbb{Z}_{\geq 0}^{r+d}$ in the following way: (i) α_R consists of the counts of the reactions in the sequence of reactions in (11), and (ii) α_D consists of the counts of the domination relations in the paths taken to construct the reaction sequences in (11).

We now show that α is balanced according to Definition 8. It is clear, first of all, that α only has support on \mathcal{R}_F and \mathcal{D}_F so that Condition 1 is satisfied. In order to show that $\alpha_R \in \ker(\Gamma)$, we note from Eq. (2) of the main text, and the definition of α_R , that

$$\mathbf{X}^{n_2-} = \mathbf{X}^{n_1-} + \Gamma \alpha_R \implies \mathbf{0} = \Gamma \alpha_R.$$

It follows that $\alpha_R \in \ker(\Gamma)$ and therefore α satisfies Condition 2 of 8. To verify Condition 3, we note that, since y^{i-} is always chosen to be the first complex exterior to \mathcal{Y} which becomes charged, the only contribution to α from the nonterminal component comes from the segments corresponding to $\mathbf{X}^{i-} \rightsquigarrow \mathbf{X}^{i+}$, i.e. the paths from y^{i-} to \mathcal{Y} . It follows that, at every complex exterior to \mathcal{Y} , the count of the reaction out is at least as great as the sum of the reactions in, and α therefore satisfies Condition 3 of Definition 8. Lemma 2 is therefore shown. Since Theorem 1 is the contrapositive of Lemma 2, we are done. \square

References

- Anderson DF (2007) A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *J Chem Phys* 127(21):214107
- Anderson DF, Kurtz TG (2011) Continuous time Markov chain models for chemical reaction networks. In: Koeppl H et al (eds) *Design and analysis of biomolecular circuits: engineering approaches to systems and synthetic biology*. Springer, Berlin, pp 3–42
- Anderson DF, Kurtz TG (2015) *Stochastic analysis of biochemical systems*. Springer, Berlin
- Anderson DF, Craciun G, Kurtz TG (2011) Product-form stationary distributions for deficiency zero chemical reaction networks. *Bull Math Biol* 72(8):1947–1970
- Anderson DF, Enciso G, Johnston MD (2014) Stochastic analysis of chemical reaction networks with absolute concentration robustness. *J R Soc Interface* 11(93):20130943
- Anderson DF, Cappelletti D, Kurtz TG (2017) Finite time distributions of stochastically modeled chemical systems with absolute concentration robustness. *SIAM J Appl Dyn Syst* 16(3):1309–1339
- Bause F, Kritzinger PS (2002) *Stochastic Petri nets: an introduction to the theory*. Vieweg Verlag 2, Aufl
- Boros B (2013) On the dependence of the existence of the positive steady states on the rate coefficients for deficiency-one mass action systems: single linkage class. *J Math Chem* 51(9):2455–2490
- Brijder R (2015) Dominant and T-invariants for Petri nets and chemical reaction networks. *Lect Notes Comput Sci* 9211:1–15
- Craciun G, Dickenstein A, Shiu A, Sturmfels B (2009) Toric dynamical systems. *J Symbolic Comput* 44(11):1551–1565
- Enciso GA (2016) Transient absolute robustness in stochastic biochemical networks. *J R Soc Interface* 13(121):20160475
- Feinberg M (1979) *Lectures on chemical reaction networks*. Unpublished written versions of lectures given at the Mathematics Research Center, University of Wisconsin. Available from: <https://crnt.osu.edu/LecturesOnReactionNetworks>

- Feinberg M (1972) Complex balancing in general kinetic systems. *Arch Ration Mech Anal* 49:187–194
- Feinberg M (1987) Chemical reaction network structure and the stability of complex isothermal reactors: I. The deficiency zero and deficiency one theorems. *Chem Eng Sci* 42(10):2229–2268
- Feinberg M (1988) Chemical reaction network structure and the stability of complex isothermal reactors: II. Multiple steady states for networks of deficiency one. *Chem Eng Sci* 43(1):1–25
- Gibson MA, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. *J Phys Chem A* 104:1876–1889
- Gillespie D (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comput Phys* 22(4):403–434
- Horn F (1972) Necessary and sufficient conditions for complex balancing in chemical kinetics. *Arch Ration Mech Anal* 49:172–186
- Horn F, Jackson R (1972) General mass action kinetics. *Arch Ration Mech Anal* 47:81–116
- Johnston MD (2014) Translated chemical reaction networks. *Bull Math Biol* 76(5):1081–1116
- Johnston MD (2016) A linear programming approach to dynamical equivalence, linear conjugacy, and the deficiency one theorem. *J Math Chem* 54(8):1612–1631
- Johnston MD (2017) A computational approach to extinction events in chemical reaction networks with discrete state spaces. Available on the ArXiv at [arXiv:1701.02014](https://arxiv.org/abs/1701.02014)
- Johnston MD, Siegel D, Szederkényi G (2012) A linear programming approach to weak reversibility and linear conjugacy of chemical reaction networks. *J Math Chem* 50(1):274–288
- Johnston MD, Pantea C, Donnell P (2016) A computational approach to persistence, permanence, and endotacticity of chemical reaction networks. *J Math Biol* 72(1):467–498
- Kurtz TG (1972) The relationship between stochastic and deterministic models for chemical reactions. *J Chem Phys* 57:2976–2978
- Lawler GF (2006) Introduction to stochastic processes. Chapman and Hall, London
- Memmi G, Roucairol G (1975) Linear algebra in net theory. In: Brauer W (ed). *Net theory and applications*, volume 84 of *Lecture notes in computer science*. Springer, pp 213–223
- Paulevé L, Craciun G, Koeppl H (2014) Dynamical properties of discrete reaction networks. *J Math Biol* 69(1):55–72
- Shinar G, Feinberg M (2010) Structural sources of robustness in biochemical reaction networks. *Science* 327(5971):1389–1391
- Szederkényi G (2010) Computing sparse and dense realizations of reaction kinetic systems. *J Math Chem* 47:551–568