Euro. Jnl of Applied Mathematics (2023), *vol.* 34, *pp.* 303–325
doi:10.1017/S0956792522000146

On the sum of chemical reactions

LINARD HOESSLY, CARSTEN WIUF, and PANQIU XIA

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark
emails: linard.hoessly@hotmail.com; wiuf@math.ku.dk; px@math.ku.dk.

(*Received 27 July 2021; revised 25 March 2022; accepted 26 April 2022; first published online 24 May 2022*)

It is standard in chemistry to represent a sequence of reactions by a single overall reaction, often called a complex reaction in contrast to an elementary reaction. Photosynthesis $6CO_2 + 6H_2O \longrightarrow C_6H_{12}O_6 + 6O_2$ is an example of such complex reaction. We introduce a mathematical operation that corresponds to summing two chemical reactions. Specifically, we define an associative and non-communicative operation on the product space $\mathbb{N}_0^n \times \mathbb{N}_0^n$ (representing the reactant and the product of a chemical reaction, respectively). The operation models the overall effect of two reactions happening in succession, one after the other. We study the algebraic properties of the operation and apply the results to stochastic reaction networks (RNs), in particular to reachability of states, and to reduction of RNs.

Keywords: Reaction network, reduction, Markov chain, graph

2020 Mathematics Subject Classification: 37N25 (Primary), 92C42, 92E20 (Secondary)

1 Introduction

Systems of chemical reactions are commonly modelled by reaction networks (RNs) [16, 12]. RNs provide a comprehensive mathematical framework for modelling systems of interacting species that is used not only in chemistry and biophysics but also in mathematical genetics [11], epidemiology [32], cellular and systems biology [44] and sociology [43]. Notable examples include the Lotka–Volterra predator–prey system [28] and the SIR model [1].

If a series of reactions occur one by one, it is natural to ask for the overall effect of the reactions, that is, the *sum* (in some sense) of the reactions. In fact, it is standard in chemistry to summarise reactions into a single overall or *complex* reaction, in contrast to elementary reactions. As an example, photosynthesis consists of a sequence of reactions, summarised into the complex reaction

$$6CO_2 + 6H_2O \longrightarrow C_6H_{12}O_6 + 6O_2$$

[38]. Graphical treatment of such sequences of reactions, that is of complex reactions, has a long history in the chemical literature, see, for example, [8, 40, 41, 37]. Here, we provide the mathematical framework for *adding* such sequences of reactions.

As an example, consider an RN describing single gene expression [42],

$$0 \rightleftharpoons R, \quad P \longrightarrow 0, \quad R \longrightarrow R + P,$$

where R denotes an mRNA molecule and P a protein. The mRNA is freely produced from a gene (the reaction $0 \longrightarrow R$), the protein is translated from the mRNA, and both protein and mRNA are degraded. Modelled as a *discrete system*, the state space is $\mathbb{N}_0^2$, pairs of integers representing the number of R and P molecules, respectively. Jumps between states are given by reaction vectors, for example, a *direct* jump from $(k, \ell) \in \mathbb{N}_0^2$ to $(k, \ell + 1)$ is possible by means of the reaction $R \longrightarrow R + P$, if the number k of R molecules is ≥ 1. If $k = 0$, then the sequence of reactions $0 \longrightarrow R$, $R \longrightarrow R + P$, $R \longrightarrow 0$ will take the system from the state $(0, \ell)$ to the state $(0, \ell + 1)$. In that case, one might describe the overall effect (the *sum*) of the reaction sequence as $0 \longrightarrow P$. As the R molecule is created in the first reaction and degraded in the third, it cancels in the sum. Using similar arguments, one can conclude that the set of reachable states from any state $(k, \ell) \in \mathbb{N}_0^2$ is all of $\mathbb{N}_0^2$.

When the number of molecules of each species (here R, P) is low (as is often the case if the system is embedded into a cellular environment), it is appropriate to consider the system as a discrete stochastic system in $\mathbb{N}_0^n$. If so, it is standard to model the changes in molecule counts by a continuous-time Markov chain [27, 3]. For example, with stochastic mass-action kinetics, the propensities for the reactions to take place have the form

$$\lambda_{y \to y'}(x) = \kappa_{y \to y'} \frac{x!}{(x-y)!} \mathbb{1}_{\{z:\, z \geq y\}}(x),$$

where $\kappa_{y \to y'}$ is a positive rate constant and $z! := \prod_{i=1}^n z_i!$ for $z \in \mathbb{N}_0^n$. A first step in the analysis of a stochastic dynamical RN is to understand the structure of the reachable sets and the irreducible classes, that is, to understand whether the system is confined to subspaces of $\mathbb{N}_0^n$, is absorbed in certain states, etc., depending on the initial state of the system.

In the following, we examine a binary sum operation on $\mathbb{N}_0^n \times \mathbb{N}_0^n$ that describes the addition of two chemical reactions, as illustrated in the single gene expression RN above. We study the operation's algebraic properties and its applications. In terms of applications, we exhibit connections to discrete RNs and reachability properties and to reductions of discrete RNs. Common to these applications is the idea of reactions happening in succession, one after the other.

Reactions often occur at different time-scales [24]. This has led to various methods for *reduction* of RNs, where *fast* reactions and/or species are eliminated (in a precise mathematical sense). These methods are generally not qualitative (or graphical) *per se*, but quantitative, and depend on whether the dynamics of the RN is stochastic [24, 6, 16] or deterministic [5, 13]. If the reactions in a sequence occur at a fast rate (i.e. with high intensity), it is natural to assume no other reactions take place before the last reaction of the sequence has occurred. Rather than describing the entire sequence of reactions, one might summarise the sequence by a single complex reaction, the overall effect. In a sense, this complex reaction is obtained by *contraction*. We define contraction of reactions through the defined sum operation and subsequently define reduced RNs. These constructs are essentially graphical in nature. We show that they relate to stochastic approaches for reduction of RNs, in particular to reduction by elimination of so-called *intermediate* and *non-interacting* species [6, 21].

Furthermore, we study graphical properties of the state space of discrete RNs concerning the operation we introduce. We show that reachability can be expressed via the sum operation, and in particular that the closure of the sum operation determines reachability.

In Section 2, we define the sum of two reactions and study the properties of the operation. In Section 3, we specialise to RNs and reachability properties. Finally, in Sections 4 and 5,

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

we study reductions of RNs. In the latter section, we draw on Section 2 and study conditions that ensure that the reduction leads to reversible (or weakly reversible, or essential) RNs.

2 Algebra on $\mathbb{N}_0^n \times \mathbb{N}_0^n$

Denote by $\mathbb{Z}$ the set of integers, and by $\mathbb{N}_0$ the set of non-negative integers. Let n be a positive integer. For $x = (x^1, \ldots, x^n)$, and $y = (y^1, \ldots, y^n)$ in $\mathbb{Z}^n$, we write $x \leq y$ if $x^i \leq y^i$ for $i = 1, \ldots, n$, and $x < y$ if $x \leq y$ and $x \neq y$. We also use the notation $x \ll y$ if $x^i < y^i$ for $i = 1, \ldots, n$. Furthermore, we let $x \vee y = (x^1 \vee y^1, \ldots, x^n \vee y^n) = (\max\{x^1, y^1\}, \ldots, \max\{x^n, y^n\})$ be the component-wise maximum, and let $x \wedge y = (x^1 \wedge y^1, \ldots, x^n \wedge y^n) = (\min\{x^1, y^1\}, \ldots, \min\{x^n, y^n\})$ be the component-wise minimum.

Definition 2.1. Let $r_1 = (y_1, y_1'), r_2 = (y_2, y_2') \in \mathbb{N}_0^n \times \mathbb{N}_0^n$. Then $r_1 \oplus r_2 = (y, y')$ is the element in $\mathbb{N}_0^n \times \mathbb{N}_0^n$ given by $y = y_1 + 0 \vee (y_2 - y_1')$ and $y' = y_2' + 0 \vee (y_1' - y_2)$.

Proposition 1. $\left(\mathbb{N}_0^n \times \mathbb{N}_0^n, \oplus\right)$ *forms a non-commutative monoid with identity* $(0, 0)$.

Proof. It is straightforward to see that $\oplus$ is a stable operation on $\mathbb{N}_0^n \times \mathbb{N}_0^n$ with $(0,0) \oplus r = r \oplus (0,0) = r$ for all $r \in \mathbb{N}_0^n \times \mathbb{N}_0^n$. Let $y_1 \neq y_2 \in \mathbb{N}_0^n$. Then

$$(y_1, y_2) \oplus (y_2, y_1) = (y_1, y_1) \neq (y_2, y_2) = (y_2, y_1) \oplus (y_1, y_2),$$

hence $\oplus$ is non-commutative. To prove associativity, we assume without loss of generality, that $n = 1$. For $n > 1$, it follows by looking at each coordinate independently. Let $r_i = \left(y_i, y_i'\right)$ for $i = 1, 2, 3$. Furthermore, let $(y, y') = (r_1 \oplus r_2) \oplus r_3$ and $(\widetilde{y}, \widetilde{y}') = r_1 \oplus (r_2 \oplus r_3)$.

Then,

$$\begin{aligned}
y &= y_1 + 0 \vee (y_2 - y_1') + 0 \vee (y_3 - y_2' - 0 \vee (y_1' - y_2)),\\
y' &= y_3' + 0 \vee (y_2' + 0 \vee (y_1' - y_2) - y_3),\\
\widetilde{y} &= y_1 + 0 \vee (y_2 + 0 \vee (y_3 - y_2') - y_1'),\\
\widetilde{y}' &= y_3' + 0 \vee (y_2' - y_3) + 0 \vee (y_1' - y_2 - 0 \vee (y_3 - y_2')).
\end{aligned}$$

By distinguishing the following three cases a) $y_2 \geq y_1'$, b) $y_2 < y_1'$ and $y_3 \geq y_2'$, and c) $y_2 < y_1'$ and $y_3 < y_2'$, it is easy to verify that $y = \widetilde{y}$. A similar argument gives $y' = \widetilde{y}'$. The proof is complete. □

The sum operation reduces to standard addition in $\mathbb{N}_0^n$ on the two axis, and it is the component-wise maximum (addition in max-plus algebras) on the diagonal. The proof of the next result is straightforward and omitted.

Proposition 2. *Let* $r_1 = (y_1, y_1'), r_2 = (y_2, y_2') \in \mathbb{N}_0^n \times \mathbb{N}_0^n$. *Then,*

(1) *If* $y_1 = y_2 = 0$, *then* $r_1 \oplus r_2 = (0, y_1' + y_2')$.

(2) *If* $y_1' = y_2' = 0$, *then* $r_1 \oplus r_2 = (y_1 + y_2, 0)$.

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

(3) *If* $y_1 = y_1'$, $y_2 = y_2'$, *then* $r_1 \oplus r_2 = (y_1 \vee y_2, y_1' \vee y_2')$.

(4) $(y_1, y_2') \leq r_1 \oplus r_2 \leq (y_1 + y_2, y_1' + y_2')$. *Furthermore, the first equality holds if and only if* $y_1' = y_2$ *and the second equality holds if and only if* $y_1' \wedge y_2 = 0$.

The next statement characterises the sum operation.

Proposition 3. *Let* $r_1 = (y_1, y_1'), r_2 = (y_2, y_2') \in \mathbb{N}_0^n \times \mathbb{N}_0^n$ *and* $r_1 \oplus r_2 = (y, y')$. *Then*

(1) $y' - y = (y_1' - y_1) + (y_2' - y_2)$,

(2) *For* $x \in \mathbb{N}_0^n$: $x \geq y$ *if and only if* $x \geq y_1$ *and* $x + (y_1' - y_1) \geq y_2$,

(3) *For* $x \in \mathbb{N}_0^n$: $x \geq y'$ *if and only if* $x \geq y_2'$ *and* $x + (y_2 - y_2') \geq y_1'$.

Oppositely, if (1) *and* (2)*, or alternatively, if* (1) *and* (3)*, are fulfilled for some operation* $\oplus$ *on* $\mathbb{N}_0^n \times \mathbb{N}_0^n$*, then it is the sum operation in Definition 2.1.*

Proof. (1) The claim follows from $y' - y = y_2' + 0 \vee (y_1' - y_2) - y_1 - 0 \vee (y_2 - y_1')$, as $0 \vee (y_1' - y_2) - 0 \vee (y_2 - y_1') = y_1' - y_2$. (2) It is a direct consequence of $y = y_1 + 0 \vee (y_2 - y_1') = y_1 \vee (y_1 + y_2 - y_1')$. (3) follows similarly.

Oppositely, assume (1) and (2) are fulfilled for some operation $\oplus$. Then, for any $(y_1, y_1') \oplus (y_2, y_2') = (y, y')$, it holds that $x \geq y$ if and only if $x \geq y_1$ and $x + y_1' - y_1 \geq y_2$, that is, $x \geq y_2 - y_1' + y_1$. This implies that $y = y_1 + 0 \vee (y_2 - y_1')$. Combining this fact with (1), we get $y' = y_2' + 0 \vee (y_1' - y_2)$. If (1) and (3) are fulfilled for some operation $\oplus$, the proof is similar. It completes the proof. □

We next introduce an equivalence relation on $\mathbb{N}_0^n \times \mathbb{N}_0^n$ under which the corresponding quotient set is a commutative group.

Definition 2.2. Let $r_1 = (y_1, y_1'), r_2 = (y_2, y_2') \in \mathbb{N}_0^n \times \mathbb{N}_0^n$. Then, r_1 and r_2 are equivalent, denoted by $r_1 \sim r_2$, if $y_1' - y_1 = y_2' - y_2$.

The following theorem follows by definition and Proposition 3(1).

Theorem 2.3. $\left(\left(\mathbb{N}_0^n \times \mathbb{N}_0^n\right) / \sim, \oplus\right)$ *forms a commutative group.*

The next proposition shows that *subtraction* might be defined on $\mathbb{N}_0^n \times \mathbb{N}_0^n$ instead of the quotient space, in some situations.

Proposition 4. *Let* $r_1 = (y_1, y_1'), \widetilde{r}_1 = (\widetilde{y}_1, \widetilde{y}_1'), r_2 = (y_2, y_2') \in \mathbb{N}_0^n \times \mathbb{N}_0^n$. *The following properties hold.*

(1) *Suppose that* $r_1 \oplus r_2 = \widetilde{r}_1 \oplus r_2$ *and* $y_2 \ll y_1'$. *Then,* $r_1 = \widetilde{r}_1$.

(2) *Suppose that* $r_2 \oplus r_1 = r_2 \oplus \widetilde{r}_1$ *and* $y_1 \ll y_2'$. *Then,* $r_1 = \widetilde{r}_1$.

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

Proof. We show property (1). The proof of property (2) is similar. If $y_2 \ll y_1'$, then by definition and the assumption that $r_1 \oplus r_2 = \widetilde{r}_1 \oplus r_2$, we get

$$y_1 = y_1 + 0 \vee (y_2 - y_1') = \widetilde{y}_1 + 0 \vee \left(y_2 - \widetilde{y}_1'\right)$$

and

$$y_2' + y_1' - y_2 = y_2' + 0 \vee (y_1' - y_2) = y_2' + 0 \vee \left(\widetilde{y}_1' - y_2\right).$$

From the second equation, we have $0 \ll y_1' - y_2 = 0 \vee \left(\widetilde{y}_1' - y_2\right)$. Hence, $y_1' - y_2 = \widetilde{y}_1' - y_2$, which implies $\widetilde{y}_1' = y_1'$. As a result of the first equality, we have $y_1 = \widetilde{y}_1$. The proof is complete. □

The condition $y_2 \ll y_1'$ in Proposition 4(1) (as well as that in (ii)) cannot be weakened, which can be seen by example.

Proposition 1 allows us to define the (non-commutative) summation of a finite sequence of elements in $\mathbb{N}_0^n \times \mathbb{N}_0^n$,

$$\oplus_{i=1}^m r_i = r_1 \oplus r_2 \oplus \cdots \oplus r_m.$$

For any $r = (y, y') \in \mathbb{N}_0^n \times \mathbb{N}_0^n$, let $r^{-1} = (y', y)$ be the *inverse* of r. Then, $r \oplus r^{-1} = (y, y)$, and $r^{-1} \oplus r = (y', y')$. The inverse is unique, and furthermore

$$\left(\oplus_{i=1}^m r_i\right)^{-1} = \oplus_{i=1}^m r_{m+1-i}^{-1}$$

for $r_1, \ldots, r_m \in \mathbb{N}_0^n \times \mathbb{N}_0^n$, $m = 1, 2, \ldots$.

Corollary 1. *If* $r_1 = (y_1, y_1'), \ldots, r_m = (y_m, y_m') \in \mathbb{N}_0^n \times \mathbb{N}_0^n$ *and* $\oplus_{i=1}^m r_i = (y, y')$, *then*

(i) $y' - y = \sum_{i=1}^m y_i' - y_i$,

(ii) *For* $x \in \mathbb{N}_0^n$: $x \geq y$ *if and only if* $x + \sum_{i=1}^k (y_i' - y_i) \geq y_{k+1}$ *for* $k = 0, 1, \ldots, m-1$.

Proof. Let $r_{(m)} = \oplus_{k=1}^m r_k$. The corollary is then a consequence of Proposition 3 and induction in m. □

A subset $A \subseteq \mathbb{N}_0^n \times \mathbb{N}_0^n$ is said to be *closed* (under $\oplus$) if for any $r_1, r_2 \in A$, $r_1 \oplus r_2 \in A$ as well. Denote by $\mathrm{cl}(A)$ the *closure* of A, that is, the collection of all $r \in \mathbb{N}_0^n \times \mathbb{N}_0^n$ that can be represented as a *finite* sum of elements in A, including the empty sum by convention, that is, $(0, 0) \in \mathrm{cl}(A)$. Thus, $\mathrm{cl}(A)$ is the smallest closed set containing $A \cup \{(0, 0)\}$, namely $\mathrm{cl}(A)$ is a subset of any closed set $A' \cup \{(0, 0)\}$ with $A \subseteq A'$.

We next introduce several notions related to *reversibility*. The concepts to be introduced are analogous to concepts in RN theory, cf. [12, 7]. In particular, the term *essential* comes from Markov chain theory, but it is also used in RN theory [7]. It is also equivalent to *recurrent*, defined in [33] (see below), which is different from recurrent in Markov chain theory.

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

Definition 2.4. Let A be a subset of $\mathbb{N}_0^n \times \mathbb{N}_0^n$. We say

(1) $r \in A$ is *reversible* in A if $r^{-1} \in A$. The set A is *reversible*, if $r \in A$ implies $r^{-1} \in A$.

(2) $r \in A$ is *weakly reversible* in A, if there exist a sequence of elements $r_1 = (y_1, y_1'), \ldots, r_m = (y_m, y_m') \in A$, such that $y_{k-1}' = y_k$ for $k = 2, \ldots, m$, and $\oplus_{i=1}^m r_i = r^{-1}$. The set A is *weakly reversible*, if for any $r \in A$, r is weakly reversible in A.

(3) A is *essential*, if $\mathrm{cl}(A)$ is reversible.

By Proposition 2(4), we have $\oplus_{i=1}^m r_i = (y_1, y_m')$ in Definition 2.4(2). Clearly, the following implications hold by definition.

Lemma 1. *Let A be a subset of $\mathbb{N}_0^n \times \mathbb{N}_0^n$. Then,*

$$A \text{ is reversible} \implies A \text{ is weakly reversible} \implies A \text{ is essential.}$$

3 RNs and reachability

In this section, we combine the algebra defined in Section 2 with RN theory and present some reachability results. By definition, an RN is a subset $\mathcal{R} \subseteq \mathbb{N}_0^n \times \mathbb{N}_0^n$, containing no elements r equivalent to (0, 0). For convenience, we allow $\mathcal{R}$ to be infinite, though this is not standard in the literature [12]. We use standard terminology for RNs and refer to an element $r = (y, y') \in \mathcal{R}$ as a *reaction*, y as the *reactant* and y' as the *product* of this reaction. The species of y are *degraded* and those of y' are *produced*. Furthermore, as is standard in the literature, we consider an RN as a graph, writing $y \longrightarrow y'$ for $(y, y') \in \mathcal{R}$, and $y \rightleftharpoons y'$ for $(y, y'), (y', y) \in \mathcal{R}$.

For $i = 1, \ldots, n$, we denote by S_i the ith unit vector in $\mathbb{N}_0^n$, such that $\mathcal{S} = \{S_1, \ldots, S_n\}$ forms a complete basis of $\mathbb{N}_0^n$. For $y = (y^1, \ldots, y^n) \in \mathbb{N}_0^n$, we thus have $y = \sum_{i=1}^n y^i S_i$. We refer to S_i as the ith *species*, and the component y^i as the *stoichiometric coefficient* of the species S_i in y.

Example 1. Consider a two-substrate mechanism [10],

$$E + A \rightleftharpoons EA, \quad EA + P \longrightarrow EQ \longrightarrow E + Q,$$

where E is an enzyme catalysing the conversion of a substrate A to another substrate Q through a third intermediate substrate P. The molecules EA and EQ are referred to as transient (or intermediate) complexes.

Using the notation introduced above, let $S_1 = (1, 0, 0, 0, 0, 0) = E$, $S_2 = (0, 1, 0, 0, 0, 0) = A$, $S_3 = (0, 0, 1, 0, 0, 0) = EA$, $S_4 = (0, 0, 0, 1, 0, 0) = P$, $S_5 = (0, 0, 0, 0, 1, 0) = EQ$ and $S_6 = (0, 0, 0, 0, 0, 1) = Q$. Then, we might write the reactions as follows:

$$(1, 1, 0, 0, 0, 0) \rightleftharpoons (0, 0, 1, 0, 0, 0, 0),$$

$$(0, 0, 1, 1, 0, 0, 0) \longrightarrow (0, 0, 0, 0, 1, 0) \longrightarrow (1, 0, 0, 0, 0, 1).$$

For example, the species E has stoichiometric coefficient 1 in $(1, 1, 0, 0, 0, 0) = S + E$.

In the stochastic theory of RNs with *finite* number of reactions, the molecule counts follow a continuous-time Markov process $\{X(t)\}_{t \geq 0}$ with state space $\mathbb{N}_0^n$. Jumps occur according to the 'firing' of reactions: The reaction $y \longrightarrow y' \in \mathcal{R}$ has transition intensity $\lambda_{y \to y'}(x)$ and when it occurs

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

the process jumps from state x to state $x+y'-y$, where $y'-y$ is the *net gain* of the reaction [3]. The Markov process satisfies the following equation:

$$P(X(t+\Delta t)=x+\xi\,|X(t)=x)=\sum_{y\to y'\in\mathcal{R}\,:\,y'-y=\xi}\lambda_{y\to y'}(x)\Delta t+o(\Delta t), \tag{3.1}$$

for $\xi\in\mathbb{Z}^n$ and some initial count $X(0)=x_0\in\mathbb{N}_0^n$. As $\mathcal{R}$ is finite, then (3.1) defines the process $\{X(t)\}_{t\geq 0}$ (provided the chain does not explode).

Generally, the transition intensities $\lambda_{y\to y'}\colon \mathbb{N}_0^n\to[0,\infty)$, for $y\longrightarrow y'\in\mathcal{R}$, are assumed to satisfy the *compatibility* condition

$$\lambda_{y\to y'}(x)>0 \quad\Longleftrightarrow\quad x\geq y, \tag{3.2}$$

or the weaker condition

$$\lambda_{y\to y'}(x)>0 \quad\Longrightarrow\quad x\geq y. \tag{3.3}$$

These have natural interpretations: A reaction $y\longrightarrow y'$ can occur (if and) only if the molecule counts are larger than or equal to y. Below, we adhere to (3.2) and note that similar statements (one-way implications) to those we derive can be achieved assuming (3.3) only.

A reaction $y\longrightarrow y'\in\mathcal{R}$ is said to be *active* on a state $x\in\mathbb{Z}^n$ if $\lambda_{y\to y'}(x)>0$, and an ordered sequence of reactions $y_1\longrightarrow y_1',\ldots,y_m\longrightarrow y_m'\in\mathcal{R}$ is said to be *active* on x if

$$\lambda_{y_k\to y_k'}\left(x+\sum_{i=1}^{k-1}y_i'-y_i\right)>0,\quad k=1,\ldots,m, \tag{3.4}$$

that is, if the sequence of reactions can happen in succession, one after the other. After each step, the molecule count is updated. In particular, an ordered sequence of reactions is active on x if and only if there is a positive probability that the Markov chain performs this sequence of reactions in the given order.

Assume the compatibility condition (3.2) holds. Then, (3.4) is equivalent to $x+\sum_{i=1}^{k-1}(y_i'-y_i)\geq y_k$ for $k=1,\ldots,m$. According to Proposition 3 and Corollary 1, this provides the following interpretation of the sum operation.

Corollary 2. *An ordered sequence of reactions $y_1\longrightarrow y_1',\ldots,y_m\longrightarrow y_m'\in\mathcal{R}$ is active on a state $x\in\mathbb{N}_0^n$, if and only if $x\geq y$, where $(y,y')=\oplus_{i=1}^m(y_i\longrightarrow y_i')$.*

For a stochastic RN, *reachability* to a state $x'\in\mathbb{N}_0^n$ or the set of reachable states from an initial state $x\in\mathbb{N}_0^n$ is often a main interest [39]. A state x *leads to* a state x' via an RN $\mathcal{R}$, or equivalently, x' is *reachable* from x if there is an active ordered sequence of $m\geq 0$ reactions $y_1\longrightarrow y_1',\ldots,y_m\longrightarrow y_m'\in\mathcal{R}$ such that $x'=x+\sum_{i=1}^m y_i'-y_i$. As a consequence of Proposition 3 and Corollary 1, we can reformulate reachability of elements in $\mathbb{N}_0^n$ via $\mathcal{R}$ as follows.

Lemma 2. *Let $\mathcal{R}$ be an RN. A state $x\in\mathbb{N}_0^n$ leads to $x'\in\mathbb{N}_0^n$ if and only if there is $(y,y')\in\mathrm{cl}(\mathcal{R})$ with $x\geq y$ and $x'=x+y'-y$; equivalently $(x,x')\geq(y,y')$ and $(x,x')\sim(y,y')$.*

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

Denote by $\mathcal{R}(x) = \{x' \in \mathbb{N}_0^n | x \text{leads to } x'\}$ the set of reachable states of $x \in \mathbb{N}_0^n$ via $\mathcal{R}$.

Corollary 3. *For two RNs $\mathcal{R}_1, \mathcal{R}_2$ on the same set of species, we have*

$$\mathrm{cl}(\mathcal{R}_1) = \mathrm{cl}(\mathcal{R}_2) \quad \Longrightarrow \quad \text{for all } x \in \mathbb{N}_0^n,\ \mathcal{R}_1(x) = \mathcal{R}_2(x).$$

Hence, having the same $\mathrm{cl}(\mathcal{R})$ for two RNs is in general stronger than having the same reachability sets for all initial states (in the latter case, the RNs are said to be structurally identical [45]).

Say a reaction $y \to y' \in \mathcal{R}$ has a *catalytic species* if there is a species S_i such that $y^i > 0, (y')^i > 0$. Then an RN with no catalytic species is an RN where no reaction has a catalytic species. For RNs without catalytic species, the previous corollary can be strengthened.

Theorem 3.1. *For two RNs $\mathcal{R}_1, \mathcal{R}_2$ on the same set of species and without catalytic species, we have*

$$\mathrm{cl}(\mathcal{R}_1) = \mathrm{cl}(\mathcal{R}_2) \quad \Longleftrightarrow \quad \text{for all } x \in \mathbb{N}_0^n,\ \mathcal{R}_1(x) = \mathcal{R}_2(x).$$

Proof. We only need to prove the right to left implication. By symmetry it is sufficient to prove $\mathrm{cl}(\mathcal{R}_1) \subseteq \mathrm{cl}(\mathcal{R}_2)$. Consider the set $B_0 = \{r \in \mathbb{N}_0^n \times \mathbb{N}_0^n | r \sim r_0\}$, $r_0 \in \mathcal{R}_1$. Any element of B_0 takes the form $r = r_0 + (y, y) \in \mathbb{N}_0^n \times \mathbb{N}_0^n$ for some $y \in \mathbb{Z}^n$. As there are no catalytic species, then $y \geq 0$ and $r \geq r_0$.

Let $r_0 = (y_0, y_0')$. By definition, y_0 leads to y_0' in $\mathcal{R}_1$. As $\mathcal{R}_1(y_0) = \mathcal{R}_2(y_0)$, then also $y_0' \in \mathcal{R}_2(y_0)$. Hence, y_0 leads to y_0' in $\mathcal{R}_2$, and by Lemma 2 there is an element $\widetilde{r} \in \mathrm{cl}(\mathcal{R}_2)$ that realises this. By definition, $\widetilde{r} \sim r_0$ and $\widetilde{r} \in B_0$, hence $\widetilde{r} \geq r_0$ from above. By Lemma 2, $r_0 \geq \widetilde{r}$, hence $\widetilde{r} = r_0$ and $r_0 \in \mathrm{cl}(\mathcal{R}_2)$. Now consider an arbitrary element $\widetilde{r} \in \mathrm{cl}(\mathcal{R}_1)$, given as $\widetilde{r} = \widetilde{r}_1 \oplus \ldots \oplus \widetilde{r}_k$ with $\widetilde{r}_i \in \mathcal{R}_1$, $i = 1, \ldots, k$. We have $\widetilde{r}_i \in \mathrm{cl}(\mathcal{R}_2)$ for $i = 1, \ldots, k$, hence also $\widetilde{r} \in \mathrm{cl}(\mathcal{R}_2)$ by the closure property. □

Finally, we characterise the property of being essential. Moreover, we prove the equivalence between essential RNs defined in Definition 2.4(3) and recurrent RNs defined in [33].

Proposition 5. *An RN $\mathcal{R}$ is essential if and only if for $x, x' \in \mathbb{N}_0^n$, if x leads to x', then x' leads to x.*

Proof. If $\mathcal{R}$ is an essential RN, then as a consequence of Lemma 2, x leads to x' whenever x' leads to x. Oppositely, assume that $\mathcal{R}$ is such that x leads to x' whenever x' leads to x for all $x, x' \in \mathbb{N}_0^n$. Then, for any $r_0 \in \mathrm{cl}(\mathcal{R})$, let $r_* = (y_*, y_*') \leq r_0$ be a minimal element of $\{r \in \mathrm{cl}(\mathcal{R}) |\ r \sim r_0\}$ (which exists by Zorn's lemma, but is not necessarily unique). Note that by Lemma 2 we have that y_* leads to y_*', hence by assumption also that y_*' leads to y_*. Then by Lemma 2 there is $\widetilde{r} \in \mathrm{cl}(\mathcal{R})$ with $\widetilde{r} \leq r_*^{-1}$ and $\widetilde{r} \sim r_*^{-1}$, which is equivalent to $\widetilde{r}^{-1} \leq r_*$ and $\widetilde{r}^{-1} \sim r_*$. Similarly, we can find $\widehat{r} \in \mathrm{cl}(\mathcal{R})$ with $\widehat{r} \leq \widetilde{r}^{-1}$ and $\widehat{r} \sim \widetilde{r}^{-1}$. Thus, we have $\widehat{r} \leq r_*$ and $\widehat{r} \sim r_*$. As r_* is chosen to be a minimal element of $\{r \in \mathrm{cl}(\mathcal{R}) |\ r \sim r_0\}$, this implies $\widehat{r} = r_*$ and thus $\widetilde{r} = r_*^{-1}$. Finally, as $\mathrm{cl}(\mathcal{R})$ is a closed set, it is enough to check the equality $r_0^{-1} = r_*^{-1} \oplus r_0 \oplus r_*^{-1}$, and so $r_0^{-1} \in \mathrm{cl}(\mathcal{R})$. The proof of Proposition 5 is complete. □

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

In particular, the result characterises and connects the property of $\mathcal{R}$ to be essential with the geometry of $\mathrm{cl}(\mathcal{R})$. Considering the isometric involution defined by the inverse r^{-1} of a reaction, we can equivalently say that $\mathcal{R}$ is essential if and only if $\mathrm{cl}(\mathcal{R})$ is symmetric with respect to the above involution.

A semi-linear set is defined as a finite union of linear sets, where a linear set is a set generated by a base vector $b \in \mathbb{Z}^n$ and *period vectors* $p_1, \ldots p_k \in \mathbb{Z}^n$ as follows [30]:

$$L(b,p) = \left\{ b + \sum_{i=1}^{k} \lambda_i p_i \middle| \lambda_1, \ldots \lambda_k \in \mathbb{N}_0 \right\}.$$

Semi-linear sets are widely studied in computer science with applications in automata theory [31], formal languages [17] and Presburger arithmetic [18], as well as in models of computation, such as Petri nets and vector addition systems [9]. In terms of RNs, the discrete dynamics of Petri nets and vector addition systems might equivalently be represented by the discrete dynamics of RNs [9]. Consequently, the reachable sets of RNs are not semi-linear in general, as this is known to be the case of Petri nets and vector addition systems [22, 46]. Here, we will be concerned with a related question, namely whether the closure $\mathrm{cl}(\mathcal{R})$ of an RN $\mathcal{R}$, considered as a subset of $\mathbb{N}_0^n \times \mathbb{N}_0^n = \mathbb{N}_0^{2n}$, is semi-linear. Simple examples suggest this might be so: if $\mathcal{R} = \{\emptyset \rightleftharpoons S_1, \emptyset \rightleftharpoons S_2, \ldots, \emptyset \rightleftharpoons S_n\}$, then $\mathrm{cl}(\mathcal{R}) = \mathbb{N}_0^{2n}$; and if $\mathcal{R} = \{S_1 \rightleftharpoons S_2\}$, then

$$\mathrm{cl}(\mathcal{R}) = \{\lambda_1(1,0,0,1) + \lambda_2(0,1,1,0) + \lambda_3(1,0,1,0) + \lambda_4(0,1,0,1) \mid \lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{N}_0\}.$$

In both cases, the closure is semi-linear, in fact linear. The second example highlights the fact that the closure is generally not contained in the linear set generated by the reactions.

Following the above discussion, we ask whether $\mathrm{cl}(\mathcal{R})$ is a semi-linear set. This is not the case, as will be seen by example. For this, we need the following lemma.

Lemma 3. *Assume $A \subseteq \mathbb{N}_0^n \times \mathbb{N}_0^n$ is semi-linear, and let $x \in \mathbb{N}_0^n$. Then $A_x = \{(a,b) \in A \mid a = x\}$ is empty or a semi-linear set as well.*

Proof. It is enough to prove it for A a linear set, as semi-linear sets are finite unions of linear sets. So assume A is linear and that A_x is non-empty. We want to show that A_x is semi-linear. Let A be given by $L(b,p)$ with base vector $b \in \mathbb{N}_0^n \times \mathbb{N}_0^n$ and non-zero period vectors $p_1, \ldots p_k \in \mathbb{N}_0^n \times \mathbb{N}_0^n$. Let $\mathrm{proj}(\cdot)$ denote the projection onto the first n-coordinates.

Without loss of generality, we let $p_1, \ldots, p_m$ be the period vectors with non-trivial projection onto the first n coordinates, that is, $\mathrm{proj}(p_i) \neq 0$, $i = 1, \ldots, m$ and $\mathrm{proj}(p_i) = 0$, $i = m+1, \ldots, k$. Then, there are finitely many vectors $\lambda = (\lambda^1, \ldots, \lambda^m) \in \mathbb{N}_0^m$, such that $\mathrm{proj}(b + \sum_{i=1}^m \lambda^i p_i) = x$, as all p_i are non-zero and non-negative. Denote this finite set by B, that is,

$$B = \left\{ b + \sum_{i=1}^{m} \lambda_i p_i \middle| \lambda \in \mathbb{N}_0^m \text{ and } \mathrm{proj}\left(b + \sum_{i=1}^{m} \lambda_i p_i \right) = x \right\}.$$

Furthermore, for the first n coordinates, if p_i has a non-zero entry whenever x in B has a zero entry, then necessarily $\lambda_i = 0$, $i = 1, \ldots, m$.

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

Finally, A_x might be written as

$$A_x = \bigcup_{c \in B} \left\{ c + \sum_{m+1}^{k} \lambda_i p_i \middle| \lambda_k, \ldots \lambda_{m+1} \in \mathbb{N}_0 \right\},$$

which is a finite union of linear sets, hence a semi-linear set. □

As it is onerous to prove that a set is not semi-linear, we consider a concrete RN. Using Lemma 3, we reduce the RN to a known example originally given for a vector addition system, which is not semi-linear [22].

Corollary 4. *There exists an RN $\mathcal{R}$ such that the closure* $\mathrm{cl}(\mathcal{R})$ *of $\mathcal{R}$ is not a semi-linear set.*

Proof. We will use that there exists a 6-dimensional vector addition system with a reachability set that is not semi-linear [22]. To conclude we translate that example to the following RN:

$$S_0 + S_2 \longrightarrow S_0 + S_1, \quad S_0 \longrightarrow S_3, \quad , S_3 + S_1 \longrightarrow S_3 + 2S_2, \quad S_3 \longrightarrow S_0 + S_4,$$

such that the reachability set of [22, Lemma 2.8], which is not semi-linear, corresponds to $\mathcal{R}(S_0 + S_2)$ (the reachability set $\mathcal{R}(x)$ with $x = S_0 + S_2$).

We construct a new RN, $\widetilde{\mathcal{R}} = \{S_5 \longrightarrow S_0 + S_2\} \cup \mathcal{R}$. Then, $\widetilde{\mathcal{R}}(S_5) = \mathrm{cl}(\widetilde{\mathcal{R}})_{S_5}$, where $\mathrm{cl}(\widetilde{\mathcal{R}})_{S_5}$ is A_x with $A = \mathrm{cl}(\widetilde{\mathcal{R}})$ and $x = S_5$ (see Lemma 3), can be written as $\{S_5\} \cup \mathcal{R}(S_0 + S_2)$. We note that the union of a finite set with a non-semi-linear set is a non-semi-linear set, hence it follows by contradiction and Lemma 3 that $\mathrm{cl}(\widetilde{\mathcal{R}})$ is not semi-linear. □

4 Reduction of RNs

In this section, we study graphical reduction of an RN to a smaller (reduced) RN in terms of the number of species, entirely based on the reactions alone and not their stochastic propensities to occur. The number of reactions of the reduced RN might be bigger or smaller than the original RN. Specifically, we provide a definition of *eliminable* species and that of a *reduced* RN, obtained by removal of a set eliminable species.

The motivation comes from studying stochastic RNs with fast-slow dynamics [6, 21]. We motivate with an example.

Example 2. A simple model of protein production is the following:

$$G \rightleftharpoons G', \quad G' \longrightarrow G' + P, \quad P \longrightarrow 0,$$

where G denotes the inactive state of a gene and G' the active state, and P is a protein produced while the gene is active [34]. The protein is subsequently degraded. One might interpret the *RN* as modelling a single polyploid cell with K copies of the gene, some of which will be in the active state, while the rest will be in the inactive state. Human cells are diploid and $K = 2$.

Assume the reactions involving the active gene in the reactant, $G' \longrightarrow G$, $G' \longrightarrow G' + P$, occur at a fast rate compared to the other two reactions. Then it is reasonable to assume that whenever a gene copy is activated, a sequence of fast reactions that eventually ends with deactivation

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

of the gene copy again occurs before a protein is degraded or another gene copy is activated. Such a sequence (including conversion of G into G') takes the form

$$G \longrightarrow G', \quad \underbrace{G' \longrightarrow G' + P, \quad \ldots, \quad G' \longrightarrow G' + P}_{k \text{ instances}}, \quad G' \longrightarrow G.$$

The net effect of the sequence is simply the sum of the reactions: $G \longrightarrow G + kP$. It appears that the active gene G' has been eliminated from the *RN* through the fast reactions $G' \longrightarrow G$, $G' \longrightarrow G' + P$.

To formalise this, let $\mathcal{U} = \{G'\}$ and $\mathcal{F} = \{G' \longrightarrow G, G' \longrightarrow G' + P\}$. Then, we say $\mathcal{U}$ is eliminable with respect to $\mathcal{F}$, resulting in the reduced *RN*,

$$\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \{P \longrightarrow 0\} \cup \{G \longrightarrow G + kP | k \in \mathbb{N}_0\}.$$

The reduced *RN* has infinitely many reactions.

In the example above, any sequence of fast reactions (those of $\mathcal{F}$) will eventually be 'terminated' by $G' \longrightarrow G$. If only $G' \longrightarrow G' + P$ is fast, while $G' \longrightarrow G$ is not, then arbitrarily many protein copies would be produced before the gene copy is deactivated again. In this case, the reduced RN does not make sense. Thus, it should be a requirement that any such sequence of fast reactions is eventually terminated. Oppositely, if only $G' \longrightarrow G$ is fast, then the reaction $G' \longrightarrow G' + P$ is essentially blocked from occurring as there will be no active gene copies. Thus, it is reasonable to remove $G \longrightarrow G + P$ from the reduced RN.

To formalise elimination and reduction, we introduce some notation. Let $\mathcal{R} \subseteq \mathbb{N}_0^n \times \mathbb{N}_0^n$ be an RN and let $\mathcal{U} \subseteq \mathcal{S}$. Furthermore, let $\mathcal{R}_\mathcal{U} \subseteq \mathcal{R}$ and $\mathcal{R}'_\mathcal{U} \subseteq \mathcal{R}$ be the subsets of reactions containing species of $\mathcal{U}$ in the *reactant* and the *product*, respectively,

$$\begin{aligned} \mathcal{R}_\mathcal{U} &= \{y \longrightarrow y' \in \mathcal{R} \mid \mathcal{U} \cap \operatorname{supp}(y) \neq \emptyset\}, \\ \mathcal{R}'_\mathcal{U} &= \{y \longrightarrow y' \in \mathcal{R} \mid \mathcal{U} \cap \operatorname{supp}(y') \neq \emptyset\}, \end{aligned} \tag{4.1}$$

where $\operatorname{supp}(x) = \{S_k | k = 1, \ldots, n, x^k > 0\}$ is the support of $x \in \mathbb{N}_0^n$. Let $\overline{\mathcal{R}} = \operatorname{cl}(\mathcal{R})$ for convenience, and denote by $\overline{\mathcal{R}}_\mathcal{U}$ and $\overline{\mathcal{R}}'_\mathcal{U}$ the collection of elements in $\overline{\mathcal{R}}$ containing species of $\mathcal{U}$ in the reactant and the product, respectively, analogously to (4.1). We also write $\mathcal{R}_0 = \mathcal{R} \setminus \left(\mathcal{R}_\mathcal{U} \cup \mathcal{R}'_\mathcal{U}\right)$ and $\overline{\mathcal{R}}_0 = \overline{\mathcal{R}} \setminus \left(\overline{\mathcal{R}}_\mathcal{U} \cup \overline{\mathcal{R}}'_\mathcal{U}\right)$. It is straightforward to see that $\overline{\mathcal{R}}_0$ is a closed set (under $\oplus$), $\overline{\mathcal{R}}_0 \supseteq \operatorname{cl}(\mathcal{R}_0)$, and in general, $\overline{\mathcal{R}}_0 \neq \operatorname{cl}(\mathcal{R}_0)$.

We proceed by defining the reduction procedure and give further examples below.

Definition 4.1. Let $\mathcal{R}$ be an RN and $\mathcal{U} \subseteq \mathcal{S}$. The species in $\mathcal{U}$ are said to be *eliminable* (and the set $\mathcal{U}$ also *eliminable*) in $\mathcal{R}$ with respect to a set of reactions $\mathcal{F} \subseteq \mathcal{R}_\mathcal{U}$, if for any $r_0 \in \mathcal{R}'_\mathcal{U}$ and any $r_1 \in \operatorname{cl}(\mathcal{F})$ such that $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_\mathcal{U}$, there exists $r_2 \in \operatorname{cl}(\mathcal{F})$ such that $r_0 \oplus r_1 \oplus r_2 \in \overline{\mathcal{R}}_0$.

The *reduced RN* associated with this elimination is $\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \mathcal{R}_0 \cup \mathcal{R}_{\mathcal{U},\mathcal{F}}$, where

$$\mathcal{R}_{\mathcal{U},\mathcal{F}} = \{r_0 \oplus r_1 \in \overline{\mathcal{R}}_0 | \, r_0 \in \mathcal{R}'_\mathcal{U}, r_1 \in \operatorname{cl}(\mathcal{F})\} \setminus \{r \in \mathbb{N}_0^n \times \mathbb{N}_0^n | r \sim (0,0)\}. \tag{4.2}$$

Recall that $0 \in \operatorname{cl}(\mathcal{F})$ (see Section 2). As a consequence, if $r_0 \oplus r_1 \in \overline{\mathcal{R}}_0$, then it follows that $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_\mathcal{U}$. In that case, if we choose $r_2 = 0 \in \operatorname{cl}(\mathcal{F})$, then $r_0 \oplus r_1 \oplus r_2 = r_0 \oplus r_1 \in \overline{\mathcal{R}}_0$. Thus,

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

when verifying eliminability of a subset of species, we only need to consider the case $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_{\mathcal{U}} \cup \overline{\mathcal{R}}_0$, that is, $r_0 \oplus r_1 \in \overline{\mathcal{R}}'_{\mathcal{U}} \setminus \overline{\mathcal{R}}_{\mathcal{U}}$.

We provide the following interpretation of eliminability. If a set $\mathcal{U}$ is eliminable, then for any reaction r_0 that has species in $\mathcal{U}$ in its product, but not in its reactant, and any finite sum of reactions from $\mathcal{F}$, that is, any $r_1 \in \mathrm{cl}(\mathcal{F})$, the following holds:

- either $r_0 \oplus r_1$ is in $\overline{\mathcal{R}}_0$, that is, it does not contain species in $\mathcal{U}$ in its reactant nor product,
- or there is a finite sum of reactions from $\mathcal{F}$, which equals r_2, and such that $r_0 \oplus r_1 \oplus r_2$ is in $\overline{R}_0$.

From Definition 4.1, it is clear that the reduced RN might be identified as a subset of $\mathbb{N}_0^{n-d} \times \mathbb{N}_0^{n-d}$, with $|\mathcal{U}| = d \leq n$.

Example 3. We return to Example 2 with $\mathcal{U} = \{G'\}$ and $\mathcal{F} = \{G' \longrightarrow G, G' \longrightarrow G' + P\}$. Choosing r_0, r_1 as in Definition 4.1, such that $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_{\mathcal{U}}$, then $r_0 = G \longrightarrow G'$, and r_1 is either the sum of k instances of the reaction $G' \longrightarrow G' + P$ or the sum of k instances of the reaction $G' \longrightarrow G' + P$ with an additional summation by $G' \longrightarrow G$. In the first case, $r_0 \oplus r_1 = G \longrightarrow G' + kP$, and in the second, $r_0 \oplus r_1 = G \longrightarrow G + kP$. In the latter case, Definition 4.1 is fulfilled by choosing $r_2 = 0 \in \mathrm{cl}(\mathcal{F})$, while in the former, the definition is fulfilled by choosing $r_2 = G' \longrightarrow G$.

We might interpret Definition 4.1 in the following way. If $r_0 \oplus r_1 = (y, y') \notin \overline{\mathcal{R}}_{\mathcal{U}}$, then y' can be produced from y alone (which contains no species of $\mathcal{U}$). If y' contains species of $\mathcal{U}$, then these can be degraded by reactions of $\mathcal{F}$. This is guaranteed by the existence of r_2. Thus, any species of $\mathcal{U}$ produced in this way can subsequently be degraded again through fast reactions.

Example 4. A more realistic model of protein production is the following [23, 25]:

$$G \rightleftharpoons G', \quad G' \longrightarrow G' + R, \quad R \longrightarrow R + P, \quad R \longrightarrow 0, \quad P \longrightarrow 0,$$

where G, G' and P are as before, and R is an intermediate molecule (the mRNA), produced by transcription of the gene. The mRNA is produced by the active gene, and each copy of the mRNA is subsequently translated into protein. Both mRNA and protein might be degraded.

Take $\mathcal{U} = \{G', R\}$ and

$$\mathcal{F} = \mathcal{R}_{\mathcal{U}} = \{G' \longrightarrow G, G' \longrightarrow G' + R, R \longrightarrow R + P, R \longrightarrow 0\}.$$

Furthermore, $\mathcal{R}'_{\mathcal{U}} = \{G \longrightarrow G', G' \longrightarrow G' + R, R \longrightarrow R + P\}$. If $r_0 \in \mathcal{R}'_{\mathcal{U}}$ and $r_1 \in \mathrm{cl}(\mathcal{F})$ are such that $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_{\mathcal{U}}$, then it must be that $r_0 = \{G \longrightarrow G'\}$. Examples of r_1, fulfilling the requirement $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_{\mathcal{U}}$, include

(i) $r_1 = G' \longrightarrow G$;, here $r_0 \oplus r_1 = G \longrightarrow G$,

(ii) $r_1 = (G' \longrightarrow G' + R) \oplus (R \longrightarrow R + P) \oplus (R \longrightarrow R + P) = G' \longrightarrow G' + R + 2P$; here $r_0 \oplus r_1 = G \longrightarrow G' + R + 2P$,

(iii) $r_1 = (G' \longrightarrow G' + R) \oplus (R \longrightarrow R + P) \oplus (R \longrightarrow 0) = G' \longrightarrow G' + P$; here $r_0 \oplus r_1 = G \longrightarrow G' + P$.

In either case, there exists $r_2 \in \mathrm{cl}(\mathcal{F})$, such that $r_0 \oplus r_1 \oplus r_2 \in \overline{\mathcal{R}}_0$: (i) $r_2 = G' \longrightarrow G$, (ii) $r_2 = (G' \longrightarrow G) \oplus (R \longrightarrow 0)$, (iii) $r_2 = G' \longrightarrow G$.

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

The set $\mathcal{U}$ is eliminable with respect to $\mathcal{F}$, resulting in the reduced *RN*,

$$\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \{P \longrightarrow 0\} \cup \{G \longrightarrow G + kP | k \in \mathbb{N}_0\},$$

which is the same *RN* as in Example 2.

We elaborate further on the properties of elimination.

Lemma 4. *Let $\mathcal{R}$ be an RN, and let $\mathcal{U} \subseteq \mathcal{S}$. If $r_0 \in \mathcal{R}'_{\mathcal{U}}$, $r_1 = \oplus_{i=1}^m r_{1i}$, $r_{1i} \in \mathcal{R}_{\mathcal{U}}$, $i = 1, \ldots, m$, such that $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_{\mathcal{U}}$, then $r_0 \in \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$ and $r_0 \oplus \left(\oplus_{i=1}^k r_{1i}\right) \in \overline{\mathcal{R}}'_{\mathcal{U}} \setminus \overline{\mathcal{R}}_{\mathcal{U}}$ for $k = 1, \ldots, m-1$. If $r_0 \oplus r_1 \in \overline{\mathcal{R}}_0$, then $r_{1m} \in \mathcal{R}_{\mathcal{U}} \setminus \mathcal{R}'_{\mathcal{U}}$ and $\oplus_{i=k}^m r_{1i} \in \overline{\mathcal{R}}_{\mathcal{U}} \setminus \overline{\mathcal{R}}'_{\mathcal{U}}$ for $k = 1, \ldots, m$.*

Proof. Let $\left(z_0, z'_0\right) = r_0$ and $\left(z_k, z'_k\right) = r_0 \oplus \left(\oplus_{i=1}^k r_{1i}\right)$, $k = 1, \ldots, m$. It follows from Proposition 2(4) that $z_0 \le z_1 \le \ldots \le z_m$. As $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_{\mathcal{U}}$, then $\operatorname{supp}(z_k) \cap \mathcal{U} = \emptyset$ for $k = 0, \ldots, m$. This implies that $r_0 = \left(z_0, z'_0\right) \notin \mathcal{R}_{\mathcal{U}}$, and thus $r_0 \in \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$. On the other hand, for $k = 1, \ldots, m$, $r_{1k} = y_k \longrightarrow y'_k \in \mathcal{R}_{\mathcal{U}}$ and by definition, $z_k = z_{k-1} + 0 \vee \left(y_k - z'_{k-1}\right)$. As $\operatorname{supp}(z_{k-1}) \cap \mathcal{U} = \operatorname{supp}(z_k) \cap \mathcal{U} = \emptyset$, we therefore necessarily have

$$\emptyset \neq \operatorname{supp}(y_k) \cap \mathcal{U} \subseteq \operatorname{supp}\left(z'_{k-1}\right) \cap \mathcal{U},$$

and hence $r_0 \oplus \left(\oplus_{i=1}^k r_{1i}\right) \in \overline{\mathcal{R}}'_{\mathcal{U}} \setminus \overline{\mathcal{R}}_{\mathcal{U}}$.

For the second part of the lemma, we have $r_0 \oplus r_1 = r_0 \oplus \left(\oplus_{i=1}^{m-1} r_{1i}\right) \oplus r_{1m}$. By Proposition 2(4), if $r_{1m} \in \mathcal{R}'_{\mathcal{U}}$, then $r_0 \oplus r_1 \in \overline{\mathcal{R}}'_{\mathcal{U}}$, which contradicts the assumption that $r_0 \oplus r_1 \in \overline{\mathcal{R}}_0$. Thus, $r_{1m} \in \mathcal{R}_{\mathcal{U}} \setminus \mathcal{R}'_{\mathcal{U}}$. Let $\left(\widetilde{z}_k, \widetilde{z}'_k\right) = \oplus_{i=m-k+1}^m r_{1i}$ for all $k = 1, \ldots, m$, and let $\left(\widetilde{z}_{m+1}, \widetilde{z}'_{m+1}\right) = r_0 \oplus \left(\oplus_{i=1}^m r_{1i}\right)$. Using Proposition 2(4) again, we get $\widetilde{z}'_1 \le \widetilde{z}'_2 \le \ldots \le \widetilde{z}'_{m+1}$. As $r_0 \oplus r_1 = \left(\widetilde{z}_{m+1}, \widetilde{z}'_{m+1}\right) \in \overline{\mathcal{R}}_0$, we have $\operatorname{supp}(z'_k) \cap \mathcal{U} = \emptyset$, and thus $\oplus_{i=m-k+1}^m r_{1i} \notin \overline{\mathcal{R}}'_{\mathcal{U}}$ for all $k = 1, \ldots, m$. Finally, for any $k \in \{1, \ldots, m\}$, as $r_{1k} \in \mathcal{R}_{\mathcal{U}}$, Proposition 2(4) implies that $\left(\widetilde{z}_k, \widetilde{z}'_k\right) = \oplus_{i=m-k+1}^m r_{1i} \in \overline{\mathcal{R}}_{\mathcal{U}}$. Therefore, $\oplus_{i=m-k+1}^m r_{1i} \in \overline{\mathcal{R}}_{\mathcal{U}} \setminus \overline{\mathcal{R}}'_{\mathcal{U}}$ for all $k = 1, \ldots, m$. It completes the proof. □

If $r \in \overline{\mathcal{R}}$ is such that $x \in \mathbb{N}_0^n$ is active on r and $\mathcal{U} \cap \operatorname{supp}(x) = \emptyset$, then $r \notin \overline{\mathcal{R}}_{\mathcal{U}}$ (cf. Corollary 2). In particular, this applies to reactions r that appear as sums of reactions $r = \oplus_{i=0}^m r_i$ with $r_0 \in \mathcal{R}'_{\mathcal{U}}$ and $r_1, \ldots, r_m \in \mathcal{F}$.

Keeping the interpretation of fast-slow dynamics in mind, let $\mathcal{F}$ consist of the fast reactions and $\mathcal{R} \setminus \mathcal{F}$ of the slow reactions. If currently in a state $x \in \mathbb{N}_0^n$ with no molecules of the species in $\mathcal{U}$, that is, $\mathcal{U} \cap \operatorname{supp}(x) = \emptyset$, and a reaction $r_0 \in \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$ occurs, producing one or more molecules of the species in $\mathcal{U}$, then usually a sequence of reactions takes place that degrades the molecules of the species $\mathcal{U}$ again. Reactions in $\mathcal{R}_{\mathcal{U}} \setminus \mathcal{F}$ have a low probability of occurring [21].

We state some trivial cases of eliminable species.

(1) If $\mathcal{U} \subseteq \mathcal{S}$ and $\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}} = \emptyset$, then $\mathcal{U}$ is eliminable with respect to any $\mathcal{F} \subseteq \mathcal{R}_{\mathcal{U}}$. In that case $\mathcal{R}_{\mathcal{U},\mathcal{F}} = \emptyset$ and $\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \mathcal{R}_0$ (cf. Lemma 4).

(2) If $\mathcal{U} = \emptyset$, then $\mathcal{R}_{\mathcal{U}} = \mathcal{R}'_{\mathcal{U}} = \emptyset$, and $\mathcal{U}$ is eliminable with respect to $\mathcal{F} = \emptyset$, and $\mathcal{R}_0 = \mathcal{R}$, $\mathcal{R}_{\mathcal{U},\mathcal{F}} = \emptyset$ and thus $\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \mathcal{R}$.

(3) If $\mathcal{U} = \mathcal{S}$, then $\mathcal{R}_0 = \emptyset$, $\mathcal{R}_{\mathcal{U}} = \mathcal{R}'_{\mathcal{U}} = \mathcal{R}$ and hence this is a special case of (i) with $\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \emptyset$.

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

If $\mathcal{U}$ is eliminable with respect to both $\mathcal{F}_1 \subseteq \mathcal{R}_{\mathcal{U}}$ and $\mathcal{F}_2 \subseteq \mathcal{R}_{\mathcal{U}}$, then $\mathcal{U}$ is not necessarily eliminable with respect to $\mathcal{F}_1 \cup \mathcal{F}_2$. The same is the case if disjoint sets $\mathcal{U}_1$ and $\mathcal{U}_2$ are eliminable with respect to $\mathcal{F}_1$ and $\mathcal{F}_2$, respectively (potentially with empty intersection), then $\mathcal{U}_1 \cup \mathcal{U}_2$ is not necessarily eliminable with respect to $\mathcal{F}_1 \cup \mathcal{F}_2$. Here, it should at least be required that $\mathcal{U}_2$ is eliminable with respect to $\mathcal{F}_2 \subseteq \left(\mathcal{R}^*_{\mathcal{U}_1,\mathcal{F}_1}\right)_{\mathcal{U}_2}$, see Proposition 7. An example of this is also given below, where the two sets of eliminable species do not appear in the same reactions; hence, the condition is trivially fulfilled.

Proposition 6. *Let $\mathcal{R}$ be an RN, and let $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2 \subseteq \mathcal{S}$ with $\mathcal{U}_1 \cap \mathcal{U}_2 = \emptyset$. Suppose that $\mathcal{U}_1$ in $\mathcal{R}$ is eliminable with respect to $\mathcal{F}_1$, that $\mathcal{U}_2$ is eliminable in $\mathcal{R}$ with respect to $\mathcal{F}_2$, and that*

$$\left(\mathcal{R}_{\mathcal{U}_1} \cup \mathcal{R}'_{\mathcal{U}_1}\right) \cap \left(\mathcal{R}_{\mathcal{U}_2} \cup \mathcal{R}'_{\mathcal{U}_2}\right) = \emptyset.$$

Then, $\mathcal{U}$ is eliminable in $\mathcal{R}$ with respect to $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$.

Proof. Let $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_{\mathcal{U}}$ with $r_0 \in \mathcal{R}'_{\mathcal{U}} = \mathcal{R}'_{\mathcal{U}_1} \cup \mathcal{R}'_{\mathcal{U}_2}$, $r_1 \in \mathrm{cl}(\mathcal{F})$. If $r_0 \oplus r_1 \notin \overline{\mathcal{R}}'_{\mathcal{U}}$, then we are done. Otherwise, suppose that $r_0 \oplus r_1 \in \overline{\mathcal{R}}'_{\mathcal{U}}$. Without loss of generality, assume that $r_0 \in \mathcal{R}'_{\mathcal{U}_1}$. Let $r_1 = \oplus_{i=1}^m r_{1i}$ with $r_{1i} \in \mathcal{F}$. As $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_{\mathcal{U}}$, then by Lemma 4, we have $r_0 \in \mathcal{R}'_{\mathcal{U}_1} \setminus \mathcal{R}_{\mathcal{U}}$ and $r_0 \oplus \left(\oplus_{i=1}^k r_{1i}\right) \in \overline{\mathcal{R}}'_{\mathcal{U}} \setminus \overline{\mathcal{R}}_{\mathcal{U}}$ for all $k = 1, \ldots, m$. Note that $\mathcal{R}'_{\mathcal{U}_1} \cap \mathcal{R}'_{\mathcal{U}_2} = \emptyset$, hence r_0 has no species of $\mathcal{U}_2$ in the product. We claim that r_{11} has no species of $\mathcal{U}_2$ in the reactant. If this is not the case, then $r_0 \oplus r_{11} \in \overline{\mathcal{R}}_{\mathcal{U}_2} \subseteq \overline{\mathcal{R}}_{\mathcal{U}}$, which contradicts the fact that $r_0 \oplus r_{11} \in \overline{\mathcal{R}}'_{\mathcal{U}} \setminus \overline{\mathcal{R}}_{\mathcal{U}}$. Thus, $r_{11} \in \mathcal{F}_1$.

Recall the assumption that $\mathcal{R}_{\mathcal{U}_1} \cap \mathcal{R}'_{\mathcal{U}_2} = \emptyset$. It follows that $r_{11} \in \mathcal{R}_{\mathcal{U}_1} \setminus \mathcal{R}'_{\mathcal{U}_2}$, and thus, by Proposition 2(4), $r_0 \oplus r_{11} \in \overline{\mathcal{R}}'_{\mathcal{U}_1} \setminus \overline{\mathcal{R}}'_{\mathcal{U}_2}$. As a result, r_{12} has no species of $\mathcal{U}_2$ in the reactant as well. This implies that $r_{12} \in \mathcal{F}_1$. Iteratively, we can show that $r_{1k} \in \mathcal{F}_1$ for all $k = 1, \ldots, m$. In other words, $r_1 = \oplus_{k=1}^m r_{1k} \in \mathrm{cl}(\mathcal{F}_1)$. Since $\mathcal{U}_1$ is eliminable with respect to $\mathcal{F}_1$, there exists $r_2 \in \mathrm{cl}(\mathcal{F}_1) \subseteq \mathrm{cl}(\mathcal{F})$ such that $r_0 \oplus r_1 \oplus r_2 \in \overline{\mathcal{R}} \setminus \left(\overline{\mathcal{R}}_{\mathcal{U}_1} \cup \overline{\mathcal{R}}'_{\mathcal{U}_1}\right)$. By assumption $\left(\mathcal{R}_{\mathcal{U}_1} \cup \mathcal{R}'_{\mathcal{U}_1}\right) \cap \left(\mathcal{R}_{\mathcal{U}_2} \cup \mathcal{R}'_{\mathcal{U}_2}\right) = \emptyset$, we have $r_0 \notin \mathcal{R}_{\mathcal{U}_2} \cup \mathcal{R}'_{\mathcal{U}_2}$ and $\mathrm{cl}(\mathcal{F}_1) \cap \left(\overline{\mathcal{R}}_{\mathcal{U}_2} \cup \overline{\mathcal{R}}'_{\mathcal{U}_2}\right) = \emptyset$. Thus, $r_0 \oplus r_1 \oplus r_2 \notin \left(\overline{\mathcal{R}}_{\mathcal{U}_2} \cup \overline{\mathcal{R}}'_{\mathcal{U}_2}\right)$, which yields that

$$r_0 \oplus r_1 \oplus r_2 \in \overline{\mathcal{R}}_0 = \overline{\mathcal{R}} \setminus \left(\overline{\mathcal{R}}_{\mathcal{U}} \cup \overline{\mathcal{R}}'_{\mathcal{U}}\right) = \overline{\mathcal{R}} \setminus \left(\overline{\mathcal{R}}_{\mathcal{U}_1} \cup \overline{\mathcal{R}}'_{\mathcal{U}_1} \cup \overline{\mathcal{R}}_{\mathcal{U}_2} \cup \overline{\mathcal{R}}'_{\mathcal{U}_2}\right).$$

This completes the proof of this lemma. □

We introduce some important classes of species that often appear in practice [14, 35, 36]. See also Example 5.

Definition 4.2. Let $\mathcal{R}$ be an RN and $\mathcal{U} \subseteq \mathcal{S}$. Then,

(1) $\mathcal{U}$ consists of *non-interacting* species, if for any two species $S_i, S_j \in \mathcal{U}$ and any reaction $y \longrightarrow y' \in \mathcal{R}$, the sum of the stoichiometric coefficients $y^i + y^j$ and $(y')^i + (y')^j$ in the reactant and the product, respectively, are at most one.

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

(2) $\mathcal{U}$ consists of *intermediate* species, if the species of $\mathcal{U}$ are non-interacting and furthermore, for $S_i \in \mathcal{U}$ and $y \longrightarrow y' \in \mathcal{R}$, whenever $y^i = 1$, then $y = S_i$, and whenever $(y')^i = 1$, then $y' = S_i$.

Example 5 (Example 1 revisited). Recall the reactions

$$E + A \rightleftharpoons EA, \quad EA + P \longrightarrow EQ \longrightarrow E + Q.$$

The set $\mathcal{U} = \{EQ\}$ consists of intermediate species and $\mathcal{U}$ is eliminable with respect to $\mathcal{F} = \{EQ \longrightarrow E + Q\}$. The reduced RN is $\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \{E + A \rightleftharpoons EA, EA + P \longrightarrow E + Q\}$. Similarly, the set $\mathcal{U} = \{EA, EQ\}$ consists of non-interacting species and $\mathcal{U}$ is eliminable with respect to $\mathcal{F} = \{EA \longrightarrow E + A, EA + P \longrightarrow EQ \longrightarrow E + Q\}$. The reduced RN is $\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \{E + A + P \longrightarrow E + Q\}$.

Lemma 5. *Let $\mathcal{R}$ be an RN and $\mathcal{U} \subseteq \mathcal{S}$ a set of non-interacting species. Furthermore, let $r_0 \in \mathcal{R}'_{\mathcal{U}}$, $r_1 = \oplus_{i=1}^m r_{1i}$, $r_{1i} \in \mathcal{R}_{\mathcal{U}}$, $i = 1, \ldots, m$, such that $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_{\mathcal{U}}$. Then,*

(1) *$r_0 \in \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$, $r_{1i} \in \mathcal{R}_{\mathcal{U}} \cap \mathcal{R}'_{\mathcal{U}}$, $i = 1, \ldots, m-1$.*

(2) *Assume $r_0 = y_0 \longrightarrow y'_0$ and $r_{1i} = y_i \longrightarrow y'_i$, $i = 1, \ldots, m$. Then, $\mathrm{supp}(y_i) \cap \mathcal{U} = \mathrm{supp}\left(y'_{i-1}\right) \cap \mathcal{U} \neq \emptyset$ for $i = 1, \ldots, m$.*

(3) *If $r_0 \oplus r_1 \in \overline{\mathcal{R}}_0$, then $r_{1m} \in \mathcal{R}_{\mathcal{U}} \setminus \mathcal{R}'_{\mathcal{U}}$.*

Oppositely, let $r_0 \in \mathcal{R}'_{\mathcal{U}}$, $r_1 = \oplus_{i=1}^m r_{1i}$, $r_{1i} \in \mathcal{R}_{\mathcal{U}}$, $i = 1, \ldots, m$. Suppose that both (1) and (2) hold. Then, $r_0 \oplus r_1 \notin \overline{\mathcal{R}}_{\mathcal{U}}$. If furthermore, $r_{1m} \in \mathcal{R}_{\mathcal{U}} \setminus \mathcal{R}'_{\mathcal{U}}$, then $r_0 \oplus r_1 \in \overline{\mathcal{R}}_0$.

Proof. The backward direction of the lemma is straightforward, so we only need to prove the forward direction. Note that Lemma 4 and (2) imply (1) and (3); hence, we are left to prove (2). Assume $m \geq 2$, as otherwise there is nothing to prove. Recall that $\mathcal{U}$ is a set of non-interacting species. Since $r_k \in \mathcal{F} \subseteq \mathcal{R}_{\mathcal{U}}$, $k = 1, \ldots, m$, then each reactant y_k contains exactly one species in $\mathcal{U}$ with stoichiometric coefficient one. Let $\left(z_k, z'_k\right) = r_0 \oplus (\oplus_{i=1}^k r_i)$, $k = 1, \ldots, m$. By repeating the proof in Lemma 4, we find that $\mathrm{supp}(y_1) \cap \mathcal{U} = \mathrm{supp}(y'_0) \cap \mathcal{U}$ and $\mathrm{supp}(y'_0 - y_1) \cap \mathcal{U} = \emptyset$. Thus, $\mathrm{supp}(z'_1) \cap \mathcal{U} = \mathrm{supp}(y'_1 + 0 \vee (y'_0 - y_1)) \cap \mathcal{U} = \mathrm{supp}(y'_1) \cap \mathcal{U}$ is a singleton. Therefore, $\mathrm{supp}(y'_1) \cap \mathcal{U} = \mathrm{supp}(z'_1) \cap \mathcal{U} = \mathrm{supp}(y_2) \cap \mathcal{U}$. The proof of this lemma can be completed by iteration. □

The conclusion of Lemma 5 is not true in *general*, see Example 9.

Proposition 7. *Let $\mathcal{R}$ be an RN and let $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2 \subseteq \mathcal{S}$ be a set of non-interacting species such that $\mathcal{U}_1 \cap \mathcal{U}_2 = \emptyset$. Furthermore, assume $\mathcal{U}_1$ is eliminable with respect to $\mathcal{F}_1 \subseteq \mathcal{R}_{\mathcal{U}_1}$ in $\mathcal{R}$, and that $\mathcal{U}_2$ is eliminable with respect to $\mathcal{F}_2 = \left(\mathcal{R}^*_{\mathcal{U}_1,\mathcal{F}_1}\right)_{\mathcal{U}_2}$ in $\mathcal{R}^*_{\mathcal{U}_1,\mathcal{F}_1}$. Then $\mathcal{U}$ is eliminable with respect to $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{R}_{\mathcal{U}_2}$ in $\mathcal{R}$.*

Proof. We make use of the following notation: $\widetilde{\mathcal{R}} = \mathcal{R}^*_{\mathcal{U}_1,\mathcal{F}_1}$, $\widehat{\mathcal{R}} = \mathrm{cl}(\widetilde{\mathcal{R}})$, $\widehat{\mathcal{R}}_0 = \widehat{\mathcal{R}} \setminus \left(\widehat{\mathcal{R}}_{\mathcal{U}_2} \cup \widehat{\mathcal{R}}'_{\mathcal{U}_2}\right)$, and $\overline{\mathcal{R}}_0 = \overline{\mathcal{R}} \setminus \left(\overline{\mathcal{R}}_{\mathcal{U}} \cup \overline{\mathcal{R}}'_{\mathcal{U}}\right)$. Then, $\widehat{\mathcal{R}}$ is a closed subset of $\overline{\mathcal{R}}$ such that for

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

any $(y,y')\in\widehat{\mathcal{R}}$, $\mathrm{supp}(y)\cap\mathcal{U}_1=\mathrm{supp}(y')\cap\mathcal{U}_1=\emptyset$. Thus, $\widehat{\mathcal{R}}\subseteq\overline{\mathcal{R}}\setminus\left(\overline{\mathcal{R}}_{\mathcal{U}_1}\cup\overline{\mathcal{R}}'_{\mathcal{U}_1}\right)$. Furthermore, by definition $\widehat{\mathcal{R}}_0$ consists of all $(y,y')\in\widehat{\mathcal{R}}$ such that $\mathrm{supp}(y)\cap\mathcal{U}_2=\mathrm{supp}(y')\cap\mathcal{U}_2=\emptyset$, thus $\widehat{\mathcal{R}}_0\cap\left(\overline{\mathcal{R}}_{\mathcal{U}_2}\cup\overline{\mathcal{R}}'_{\mathcal{U}_2}\right)=\emptyset$. It follows that

$$\widehat{\mathcal{R}}_0\subseteq\overline{\mathcal{R}}\setminus\left(\overline{\mathcal{R}}_{\mathcal{U}_1}\cup\overline{\mathcal{R}}'_{\mathcal{U}_1}\cup\overline{\mathcal{R}}_{\mathcal{U}_2}\cup\overline{\mathcal{R}}'_{\mathcal{U}_2}\right)=\overline{\mathcal{R}}_0.$$

By definition, we need to verify that for any $r_0\in\mathcal{R}'_{\mathcal{U}}$, $r_1=\oplus_{i=1}^m r_{1i}$, $r_{1i}\in\mathcal{F}$, with $r_0\oplus r_1\notin\overline{\mathcal{R}}_{\mathcal{U}}$, either $r_0\oplus r_1\in\overline{\mathcal{R}}_0$, or there exists $r_2\in\mathrm{cl}(\mathcal{F})$ such that

$$r_0\oplus r_1\oplus r_2\in\overline{\mathcal{R}}_0. \tag{4.3}$$

Before proving this property, we show that $\mathrm{cl}(\mathcal{F}_1)\subseteq\mathrm{cl}(\mathcal{F})$ and $\mathrm{cl}(\mathcal{F}_2)\subseteq\mathrm{cl}(\mathcal{F})$. The first inclusion is trivial. Now we prove the second one. Let $r\in\mathcal{F}_2$. Then, by definition either $r\in\mathcal{F}_2\cap\left[\mathcal{R}\setminus\left(\mathcal{R}_{\mathcal{U}_1}\cup\mathcal{R}'_{\mathcal{U}_1}\right)\right]$ or $r\in\mathcal{F}_2\cap\mathcal{R}_{\mathcal{U}_1,\mathcal{F}_1}$. In the former case, we have $r\in\mathcal{F}_2\cap\mathcal{R}_{\mathcal{U}_2}\subseteq\mathcal{F}$. Thus, it suffice to consider the second case, for which, we can write $r=\widehat{r}_0\oplus\widehat{r}_1\in\mathcal{R}_{\mathcal{U}_1,\mathcal{F}_1}$ with $\widehat{r}_0\in\mathcal{R}'_{\mathcal{U}_1}$ and $\widehat{r}_1=\oplus_{k=1}^m\widehat{r}_{1k}$, $\widehat{r}_{1k}=\widehat{y}_{1k}\longrightarrow\widehat{y}'_{1k}\in\mathcal{F}_1\subseteq\mathcal{R}_{\mathcal{U}_1}$. Thus, $\mathrm{supp}(y_{1k})\cap\mathcal{U}_1\neq\emptyset$, and by the non-interacting property, it holds that $\mathrm{supp}(y_k)\cap\mathcal{U}_2=\emptyset$ for all $k=1,\ldots,m$. Therefore, due to Proposition 2(4), we have $\widehat{r}_1\in\overline{\mathcal{R}}_{\mathcal{U}_1}\setminus\overline{\mathcal{R}}_{\mathcal{U}_2}$. Hence, by Proposition 2(4) again and because $\widehat{r}_0\oplus\widehat{r}_1\in\mathcal{F}_2\subseteq\overline{\mathcal{R}}_{\mathcal{U}_2}$ has a non-interacting species in $\mathcal{U}_2$ in the reactant, $\widehat{r}_0\in\mathcal{R}_{\mathcal{U}_2}\subseteq\mathcal{F}$. Thus, $r\in\mathrm{cl}(\mathcal{F})$.

Suppose that $r_0\oplus r_1\in\overline{\mathcal{R}}'_{\mathcal{U}}\setminus\overline{\mathcal{R}}_{\mathcal{U}}$. Next, we show the existence of an $r_2\in\mathrm{cl}(\mathcal{F})$, such that (4.3) holds. Recall $r_1=\oplus_{i=1}^m r_{1i}$, $r_{1i}\in\mathcal{F}$ such that $r_0\oplus r_1=r_0\oplus\left(\oplus_{i=1}^m r_{1i}\right)\in\overline{\mathcal{R}}'_{\mathcal{U}}\setminus\overline{\mathcal{R}}_{\mathcal{U}}$. We claim that $r_{1m}\in\mathcal{R}'_{\mathcal{U}}$. Otherwise, assume $r_{1m}\in\mathcal{R}_{\mathcal{U}}\setminus\mathcal{R}'_{\mathcal{U}}$. By the opposite part of Lemma 5, we have $r_0\oplus r_1\in\overline{\mathcal{R}}_0$, which contradicts the assumption that $r_0\oplus r_1\in\overline{\mathcal{R}}'_{\mathcal{U}}\setminus\overline{\mathcal{R}}_{\mathcal{U}}$. Thus, we have $r_{1m}\in\mathcal{R}_{\mathcal{U}}\cap\mathcal{R}'_{\mathcal{U}}$. Suppose that $r_{1m}\in\mathcal{R}_{\mathcal{U}_1}\cap\mathcal{R}'_{\mathcal{U}_1}$. Let j be the largest index strictly smaller than m such that $r_{1j}\notin\mathcal{R}_{\mathcal{U}_1}$ (with $r_{10}=r_0$). If $j=0$, then $r_{1i}\in(\mathcal{R}_{\mathcal{U}_1}\cap\mathcal{R}'_{\mathcal{U}_1})\cap\mathcal{F}\subseteq\mathcal{F}_1$, $i=1,\ldots,m$ and, by Lemma 4, $r_0\in\mathcal{R}'_{\mathcal{U}_1}\setminus\mathcal{R}_{\mathcal{U}_1}$. As $\mathcal{U}_1$ is eliminable with respect to $\mathcal{F}_1$, there exists $\widehat{r}_2\in\mathrm{cl}(\mathcal{F}_1)$ such that $\widehat{r}=r_0\oplus r_1\oplus\widehat{r}_2\in\widetilde{\mathcal{R}}$. If $\widehat{r}\notin\widetilde{\mathcal{R}}'_{\mathcal{U}_2}$, then (4.3) holds with $r_2=\widehat{r}_2$. Otherwise, $\widehat{r}\in\widetilde{\mathcal{R}}'_{\mathcal{U}_2}\setminus\widetilde{\mathcal{R}}_{\mathcal{U}_2}$. Since $\mathcal{U}_2$ is eliminable in $\widetilde{\mathcal{R}}$, with respect to $\mathcal{F}_2$, there exists $\widehat{r}_3\in\mathrm{cl}(\mathcal{F}_2)\subseteq\mathrm{cl}(\mathcal{F})$ such that $r_0\oplus r_1\oplus\widehat{r}_2\oplus\widehat{r}_3\in\widehat{\mathcal{R}}_0$. Thus, we get (4.3) with $r_2=\widehat{r}_2\oplus\widehat{r}_3$.

On the other hand, if $j>0$, then $r_{1j}\in\mathcal{R}_{\mathcal{U}_2}\cap\mathcal{R}'_{\mathcal{U}_1}$ and $r_{1(j+1)},\ldots,r_{1m}\in\mathcal{F}_1$. Thus, by eliminability of $\mathcal{U}_1$ in $\mathcal{R}$ with respect to $\mathcal{F}_1$, there exists $\widehat{r}_2\in\mathrm{cl}(\mathcal{F}_1)$ such that $r_{1j}\oplus\cdots\oplus r_{1m}\oplus\widehat{r}_2\in\widetilde{\mathcal{R}}_{\mathcal{U}_2}$. Let j' be the largest index strictly smaller than j such that $r_{1j'}\notin\mathcal{R}_{\mathcal{U}_2}$. If $j'=0$, then $r_0\in\mathcal{R}'_{\mathcal{U}_2}\setminus\mathcal{R}_{\mathcal{U}}\subseteq\widetilde{\mathcal{R}}'_{\mathcal{U}_2}\setminus\widetilde{\mathcal{R}}_{\mathcal{U}_2}$. Otherwise, $r_{1j'}\in\mathcal{R}_{\mathcal{U}_1}\cap\mathcal{R}'_{\mathcal{U}_2}$ and $r_{1(j'+1)},\ldots,r_{1j}\in\mathcal{R}_{\mathcal{U}_2}\cap\mathcal{R}_{\mathcal{U}'_2}\subseteq\widetilde{\mathcal{R}}_{\mathcal{U}_2}=\mathcal{F}_2$. Let j" be the largest index strictly smaller than j' such that $r_{1j''}\notin\mathcal{R}_{\mathcal{U}_1}$. By using the opposite part of Lemma 5, we see that $r_{1j''}\oplus\cdots\oplus r_{1j'}\in\widetilde{\mathcal{R}}_{\mathcal{U}_2}$. By repeating the same argument, we find that $r_0,r_{11},\cdots,r_{1(j''-1)}$ can be divided into ordered groups such that the sum of the reactions in each group, except the first group, is either in $(\mathcal{R}_{\mathcal{U}_1,\mathcal{F}_1})_{\mathcal{U}_2}$ or $\mathcal{R}_{\mathcal{U}_2}\cap\mathcal{R}'_{\mathcal{U}_2}$, which are both in $\widetilde{\mathcal{R}}_{\mathcal{U}_2}$, and the sum of the reactions in the first group is in $\widetilde{\mathcal{R}}'_{\mathcal{U}_2}\setminus\widetilde{\mathcal{R}}_{\mathcal{U}_2}$. Hence, the existence of r_2 follows from eliminability of $\mathcal{U}_2$ with respect to $\mathcal{F}_2=\widetilde{\mathcal{R}}_{\mathcal{U}_2}$ in $\widetilde{\mathcal{R}}$.

The other cases when r_{1m} is in $\mathcal{R}_{\mathcal{U}_1}\cap\mathcal{R}'_{\mathcal{U}_2}$, $\mathcal{R}_{\mathcal{U}_2}\cap\mathcal{R}'_{\mathcal{U}_1}$ or $\mathcal{R}_{\mathcal{U}_2}\cap\mathcal{R}'_{\mathcal{U}_2}$ are essentially proved by the same means as above. The proof of the proposition is complete. □

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

It is not sufficient that $\mathcal{U}_2$ is eliminable with respect to $\mathcal{R}_{\mathcal{U}_2}$ in $\mathcal{R}$. For example, consider the RN $\mathcal{R} = \{S_1 \longrightarrow U_1 \rightleftharpoons U_2\}$. Then $\mathcal{U}_1 = \{U_1\}$ is eliminable with respect to $\mathcal{R}_{\mathcal{U}_1}$, and $\mathcal{U}_2 = \{U_2\}$ is eliminable with respect to $\mathcal{R}_{\mathcal{U}_2}$. However, $\mathcal{U}_1 \cup \mathcal{U}_2$ is not eliminable with respect to $\mathcal{R}_{\mathcal{U}_1 \cup \mathcal{U}_2}$.

We further note that several approaches to reductions of RNs derived from a deterministic dynamical perspective have been studied both in terms of slow-fast dynamics [26, 13] and in the context of steady states [19, 15, 14, 35, 29] for intermediates and non-interacting species in general. In the case of intermediates, our reduced RN agrees with the one suggested in [15]. However, for non-interacting species the reduced RN we obtain differs from that of [13, 35]. This is a consequence of the discrete nature of the state space in our case compared to the continuous state space for deterministic reaction systems.

5 Reversibility analysis for reduced RNs

Reversibility (weak reversibility, essentiality) is an important property for an RN and often implies strong properties on the dynamics, irrespectively whether the RN is modelled deterministically or stochastically [2, 7, 12, 3, 4]. Therefore, we are interested in finding criteria for a reduced RN to be reversible (weakly reversible, essential), provided the original RN is. However, in general, this appears to be a challenging problem. Here, we provide sufficient conditions for a reduced RN to be (weakly) reversible under the assumption that the eliminable species are non-interacting species.

For a set $A \subseteq \mathbb{N}_0^n \times \mathbb{N}_0^n$, let $A^{-1} = \{r^{-1} | r \in A\}$.

Theorem 5.1. *Let $\mathcal{R}$ be an RN and $\mathcal{U} \subseteq \mathcal{S}$ a set of non-interacting species. Assume $\mathcal{U}$ is eliminable with respect to $\mathcal{F} \subseteq \mathcal{R}_{\mathcal{U}}$, as in Definition 4.1, and define the condition*

$$(*) \qquad \left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right)^{-1} = \mathcal{F} \setminus \mathcal{R}'_{\mathcal{U}} \quad \text{and} \quad \mathcal{F} \cap \mathcal{R}'_{\mathcal{U}} \text{ is essential.}$$

Then,

(1) *If $\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right) \cup \mathcal{F}$ is reversible, then $(*)$ holds.*

(2) *$\mathcal{R}_{\mathcal{U},\mathcal{F}}$ is reversible if $(*)$ holds.*

(3) *$\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$ is (weakly) reversible if $\mathcal{R}_0$ is (weakly) reversible and $(*)$ holds.*

(4) *$\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$ is weakly reversible if there exists $\mathcal{F}_0 \subseteq \mathrm{cl}(\mathcal{R})$ such that $(\mathcal{R} \setminus \mathcal{R}_{\mathcal{U}}) \cup \mathcal{F}_0$ is weakly reversible and $(*)$ holds.*

Proof. (1) Firstly, note that $\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right) \cup \mathcal{F}$ can be decomposed into three disjoint sets $\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$, $\mathcal{F} \setminus \mathcal{R}'_{\mathcal{U}} \subseteq \mathcal{R}_{\mathcal{U}} \setminus \mathcal{R}'_{\mathcal{U}}$ and $\mathcal{F} \cap \mathcal{R}'_{\mathcal{U}} \subseteq \mathcal{R}_{\mathcal{U}} \cap \mathcal{R}'_{\mathcal{U}}$. Since for any $(y, y') \in \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$, $\mathrm{supp}(y) \cap \mathcal{U} = \emptyset$ and $\mathrm{supp}(y') \cap \mathcal{U} \neq \emptyset$, it follows that $(y', y) \notin \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$. For the same reason, $(y', y) \notin \mathcal{F} \cap \mathcal{R}'_{\mathcal{U}}$, where $\mathcal{F} \cap \mathcal{R}'_{\mathcal{U}} \subseteq \mathcal{R}_{\mathcal{U}} \cap \mathcal{R}'_{\mathcal{U}}$. Thus, we have

$$\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right)^{-1} \cap \left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right) = \left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right)^{-1} \cap \left(\mathcal{F} \cap \mathcal{R}'_{\mathcal{U}}\right) = \emptyset. \tag{5.1}$$

By reversibility of $\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right) \cup \mathcal{F}$, $\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right)^{-1} \subseteq \left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right) \cup \mathcal{F} = \left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right) \cup (\mathcal{F} \setminus \mathcal{R}'_{\mathcal{U}}) \cup \left(\mathcal{F} \cap \mathcal{R}'_{\mathcal{U}}\right)$. Combining this fact with (5.1), we have $\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right)^{-1} \subseteq \mathcal{F} \setminus \mathcal{R}'_{\mathcal{U}}$. Similarly, it holds that $(\mathcal{F} \setminus \mathcal{R}'_{\mathcal{U}})^{-1} \subseteq \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$, which, together with $\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right)^{-1} \subseteq \mathcal{F} \setminus \mathcal{R}'_{\mathcal{U}}$, implies

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

$\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right)^{-1} = \mathcal{F} \setminus \mathcal{R}'_{\mathcal{U}}$. For the same reason, we can show that $\left(\mathcal{F} \cap \mathcal{R}'_{\mathcal{U}}\right)^{-1} \subseteq \mathcal{F} \cap \mathcal{R}'_{\mathcal{U}}$ holds. Hence, $\mathcal{F} \cap \mathcal{R}'_{\mathcal{U}}$ is essential. In other words, $(*)$ is true and the proof is complete.

(2) Let $r_0 \oplus r_1 \in \mathcal{R}_{\mathcal{U},\mathcal{F}}$, where r_0, r_1 are as in Definition 4.1, equation (4.2). Furthermore, there exists $r_{11}, \ldots, r_{1m} \in \mathcal{F}$, such that $r_1 = \oplus_{i=1}^{m} r_{1i}$. By Lemma 5, $r_0 \in \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$, $r_{1m} \in \left(\mathcal{R}_{\mathcal{U}} \setminus \mathcal{R}'_{\mathcal{U}}\right) \cap \mathcal{F} = \mathcal{F} \setminus \mathcal{R}'_{\mathcal{U}}$ and $\{r_{11}, \ldots, r_{1(m-1)}\} \subseteq \mathcal{F} \cap \mathcal{R}'_{\mathcal{U}}$, assuming $m \geq 2$. Therefore, under condition $(*)$, we know that $r_0^{-1} \in \mathcal{F} \cap \left(\mathcal{R}_{\mathcal{U}} \setminus \mathcal{R}'_{\mathcal{U}}\right)$, $r'_0 := r_{1m}^{-1} \in \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$, and

$$\left(\oplus_{i=1}^{m-1} r_{1i}\right)^{-1} \in \mathrm{cl}(\mathcal{F} \cap \mathcal{R}'_{\mathcal{U}}) \subseteq \mathrm{cl}(\mathcal{F}).$$

Therefore, $r'_1 := \left(\oplus_{i=1}^{m-1} r_{1i}\right)^{-1} \oplus r_0^{-1} \in \mathrm{cl}(\mathcal{F})$ and thus $(r_0 \oplus r_1)^{-1} = r'_1 \oplus r'_0 \in \mathcal{R}_{\mathcal{U},\mathcal{F}}$. This proves property (2).

(3) It is a direct consequence of (2) and the definition of $\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$.

(4) It suffices to show that every $r = y \longrightarrow y' \in \mathcal{R}_0$ is weakly reversible in $\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$. Note that $\mathcal{R}_0 \subseteq \mathcal{R} \setminus \mathcal{R}_{\mathcal{U}} \subseteq (\mathcal{R} \setminus \mathcal{R}_{\mathcal{U}}) \cup \mathcal{F}_0$. Thus, by assumption, there exist reactions $y' \longrightarrow y_1, y_1 \longrightarrow y_2, \ldots, y_m \longrightarrow y \in (\mathcal{R} \setminus \mathcal{R}_{\mathcal{U}}) \cup \mathcal{F}_0$. If for $k = 1, \ldots, m$, $\mathrm{supp}(y_k) \cap \mathcal{U} = \emptyset$, then r is weakly reversible in $\mathcal{R}_0$ and thus in $\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$. Otherwise, let $i = \min\{k | \mathrm{supp}(y_k) \cap \mathcal{U} \neq \emptyset)\}$. Then

$$\{y' \longrightarrow y_1, y_1 \longrightarrow y_2, \ldots, y_{i-2} \longrightarrow y_{i-1}\} \subseteq \mathcal{R}_0 \subseteq \mathcal{R}^*_{\mathcal{U},\mathcal{F}},$$

and $y_{i-1} \longrightarrow y_i \in \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$ (with $y_0 = y'$). Let $j = \min\{k > i | \mathrm{supp}(y_k) \cap \mathcal{U} = \emptyset)\}$. Then,

$$\{y_i \longrightarrow y_{i+1}, \ldots, y_{j-1} \longrightarrow y_j\} \subseteq \mathcal{F}_0.$$

Therefore, $(y_i, y_j) = \oplus_{\ell=i}^{j} (y_{\ell-1} \longrightarrow y_\ell) \in \mathrm{cl}(\mathcal{F})$, which implies either $(y_{i-1}, y_j) = y_{i-1} \longrightarrow y_i \oplus (y_i, y_j) \in \mathcal{R}_{\mathcal{U},\mathcal{F}}$ or $\sim (0, 0)$, see Lemma 5. Repeating this process, we can find a sequence of reactions $r'_1, \ldots, r'_p$ in the reduced RN $\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$ (after removing elements equivalent to $(0, 0)$) such that the product of r'_k coincides with the reactant of r'_{k+1} for $k = 1, \ldots, p-1$, and $\oplus_{k=1}^{p} r'_k = y' \longrightarrow y$. The proof of property (4) is complete. □

We present some examples that show the limitations of Theorem 5.1.

Example 6. Consider the *RN*

$$\mathcal{R} = \{S_1 \longrightarrow U_1, U_1 \longrightarrow S_2, S_2 \longrightarrow S_1\}$$

with $\mathcal{U} = \{U_1\}$. Let $\mathcal{F} = \{U_1 \longrightarrow S_2\}$. Then, the reduced network $\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \{S_1 \rightleftharpoons S_2\}$ is reversible. However,

(1) $\mathcal{R}_0 = \{S_2 \longrightarrow S_3 + S_4\}$ is not reversible,

(2) $\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right)^{-1} = \{U_1 \longrightarrow S_1\} \neq \mathcal{F} \setminus \mathcal{R}'_{\mathcal{U}} = \{U_1 \longrightarrow S_2\}$.

Example 7. Concerning Theorem (5.1)(4), consider the *RN*

$$\mathcal{R} = \{S_1 + S_2 \longrightarrow S_3 + S_4, S_3 \longrightarrow U_1, S_4 + U_1 \longrightarrow S_1 + U_2, U_2 \longrightarrow S_2\}$$

with $\mathcal{U} = \{U_1, U_2\}$. Let $\mathcal{F} = \{S_4 + U_1 \longrightarrow S_1 + U_2, U_2 \longrightarrow S_2\}$. Then, the reduced network $\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \{S_1 + S_2 \rightleftharpoons S_3 + S_4\}$ is reversible. However,

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

(1) $\mathcal{R}_0 = \{S_1 + S_2 \longrightarrow S_3 + S_4\}$ is not reversible.

(2) $\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right)^{-1} = \{U_1 \longrightarrow S_3\} \neq \mathcal{F} \cap \left(\mathcal{R}_{\mathcal{U}} \setminus \mathcal{R}'_{\mathcal{U}}\right) = \{U_2 \longrightarrow S_2\}$.

(3) There does not exist a subset $\mathcal{F}_0 \subseteq \mathrm{cl}(\mathcal{F})$ such that $(\mathcal{R} \setminus \mathcal{R}_{\mathcal{U}}) \cup \mathcal{F}_0$ is essential, because $U_1 \longrightarrow S_3 \in (\mathcal{R} \setminus \mathcal{R}_{\mathcal{U}})^{-1}$ cannot be represented as a sum of reactions in $(\mathcal{R} \setminus \mathcal{R}_{\mathcal{U}}) \cup \mathrm{cl}(\mathcal{F})$.

Therefore, Example 6 and Example 7 imply that the conditions provided in Theorem 5.1 are not necessary conditions for (weakly) reversibility of the reduced RN. The next example shows that weak reversibility of $\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right) \cup \mathcal{F}$ in the case of non-interacting species does not ensure weak reversibility of the reduced network, implying reversibility in Theorem 5.1(1) cannot be replaced by weak reversibility and assumption $(*)$ cannot be removed in Theorem 5.1(4).

Example 8. Consider the *RN*

$$\mathcal{R} = \{S_1 \longrightarrow U_1 \longrightarrow S_2 \longrightarrow U_2 \longrightarrow S_1, S_3 + U_2 \rightleftharpoons S_4\},$$

with $\mathcal{U} = \{U_1, U_2\}$, $\mathcal{F} = \mathcal{R}_{\mathcal{U}}$. Then, $(\mathcal{R} \setminus \mathcal{R}_{\mathcal{U}}) \cup \mathcal{F} = \left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right) \cup \mathcal{F} = \mathcal{R}$ is weakly reversible, but $\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \{S_1 \rightleftharpoons S_2, S_2 + S_3 \longrightarrow S_4 \longrightarrow S_3 + S_1\}$ is not weakly reversible.

The example below shows that Theorem 5.1 is not true beyond non-interacting species.

Example 9. Consider the *RN* given by

$$\mathcal{R} = \{S_1 \rightleftharpoons U_1 + U_2, S_2 \rightleftharpoons U_1, S_3 \rightleftharpoons U_2\}$$

with $\mathcal{U} = \{U_1, U_2\}$, and let $\mathcal{F} = \mathcal{R}_{\mathcal{U}}$. Then, $(\mathcal{R} \setminus \mathcal{R}_{\mathcal{U}}) \cup \mathcal{F} = \mathcal{R}$ is reversible, $\left(\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}\right)^{-1} = \mathcal{F} \cap \left(\mathcal{R}_{\mathcal{U}} \setminus \mathcal{R}'_{\mathcal{U}}\right) = \mathcal{F}$ and $\mathcal{F} \cap \mathcal{R}'_{\mathcal{U}} = \mathcal{R}_0 = \emptyset$. In particular (1)–(4) with $\mathcal{F}_0 = \mathcal{F}$ of Theorem 5.1 are all fulfilled, but $\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \{S_1 \longrightarrow S_2 + S_3\}$ is not weakly reversible.

The last theorem of this section concerns reachability of the original and reduced RNs.

Theorem 5.2. *Let $\mathcal{R}$ be an RN and assume $\mathcal{U} \subseteq \mathcal{S}$ is eliminable with respect to $\mathcal{F} \subseteq \mathcal{R}_{\mathcal{U}}$, as in Definition 4.1. Let $x, x' \in \mathbb{N}_0^n$.*

(1) *If x leads to x' via $\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$, then x leads to x'.*

(2) *Reversely, suppose that $\mathcal{U}$ consists of intermediate species and $\mathcal{F} = \mathcal{R}_{\mathcal{U}}$. Assume* $(\mathrm{supp}(x) \cup \mathrm{supp}(x')) \cap \mathcal{U} = \emptyset$. *Then if x leads to x' via $\mathcal{R}$, then x leads also to x' via $\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$.*

Proof. (1) It follows directly from the definition of the reduced RN.

(2) Suppose x leads to x' in $\mathcal{R}$ and $(\mathrm{supp}(x) \cup \mathrm{supp}(x')) \cap \mathcal{U} = \emptyset$. Then by Lemma 2 there are reactions $r_1 \ldots, r_m \in \mathcal{R}$ (possibly with repetitions) such that $\oplus_{i=1}^m r_i \leq (x, x')$ and $\oplus_{i=1}^m r_k \sim (x, x')$. Without loss of generality, assume $\oplus_{i=1}^m r_i = (x, x')$. If this is not the case, then we proceed with $(z, z') = \oplus_{i=1}^m r_i$, rather than (x, x'), and show that $(z, z') \in \mathrm{cl}\left(\mathcal{R}^*_{\mathcal{U},\mathcal{F}}\right)$. This subsequently implies that x leads to x' via $\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$ as $\oplus_{i=1}^m r_k \sim (x, x')$.

If $r_1, \ldots, r_m \in \mathcal{R}_0$, then $\oplus_{i=1}^m r_i \in \mathrm{cl}(\mathcal{R}^*_{\mathcal{U},\mathcal{F}})$, and we are done. Otherwise, since $\mathrm{supp}(x) \cap \mathcal{U} = \emptyset$, by Lemma 5, the reaction in $\{r_1, \ldots, r_m\} \cap \left(\mathcal{R}_{\mathcal{U}} \cup \mathcal{R}'_{\mathcal{U}}\right)$ with the smallest index belongs

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

to $\mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$. Without loss of generality, assume this reaction is $r_1 = x_1 \longrightarrow u_1$, where $u_1 \in \mathcal{U}$ ($\mathcal{U}$ consists of intermediate species). Otherwise, if r_k is the first one, then $r_1, \dots, r_{k-1} \in \mathcal{R}_0 \subseteq \mathcal{R}^*_{\mathcal{U},\mathcal{F}}$, and we might define $r'_1 = r_k, r'_2 = r_{k+1}, \dots, r'_{m-k+1} = r_m$. Proceeding with the same argument as below, one can show that $\oplus_{i=1}^{m-k+1} r'_i \in \mathrm{cl}\big(\mathcal{R}^*_{\mathcal{U},\mathcal{F}}\big)$, and thus $r_1 \oplus \cdots \oplus r_{k-1} \oplus r'_1 \oplus \cdots \oplus r'_{m-k+1} \in \mathrm{cl}\big(\mathcal{R}^*_{\mathcal{U},\mathcal{F}}\big)$ as well. Hence, we take $k = 1$.

Since $\mathrm{supp}(x') \cap \mathcal{U} = \emptyset$, then there exists $k \in \{2, \dots, m\}$, such that u_1 is the reactant of r_k, but not that of $r_2, \dots, r_{k-1}$. Let $r_{2:k-1} = (x_{2:k-1}, x'_{2:k-1}) = \oplus_{i=2}^{k-1} r_i$. We claim that

$$r_1 \oplus r_k \oplus r_{2:k-1} \leq \oplus_{i=1}^{k} r_i \text{and } r_1 \oplus r_k \oplus r_{2:k-1} \sim \oplus_{i=1}^{k} r_i. \tag{5.2}$$

The equivalence in (5.2) is a consequence of Theorem 2.3. It suffices to show the inequality. Let $r_k = u_1 \longrightarrow x_2$, then $r_1 \oplus r_k = (x_1, x_2)$ and thus

$$r_1 \oplus r_k \oplus r_{2:k-1} = \big(x_1 + 0 \vee (x_{2:k-1} - x_2), x'_{2:k-1} + 0 \vee (x_2 - x_{2:k-1})\big)\,.$$

On the other hand, by the choice of r_1 and r_k, we have

$$\begin{aligned} \oplus_{i=1}^{k} r_i = r_1 \oplus r_{2:k+1} \oplus r_k &= \big(x_1 + x_{2:k-1}, u_1 + x'_{2:k-1}\big) \oplus (u_1, x_2) \\ &= \big(x_1 + x_{2:k-1}, x'_{2:k-1} + x_2\big)\,. \end{aligned}$$

This proves conclusion (5.2). Note that $r_k = u_1 \longrightarrow x_2$ implies that either $x_2 = u_2 \in \mathcal{U}$ or $\mathrm{supp}(x_2) \cap \mathcal{U} = \emptyset$. Thus, the procedure can be repeated to obtain $r_{\sigma(1)}, \dots, r_{\sigma(m)}$, where σ is a permutation of $\{1, \dots, m\}$, such that

$$\oplus_{i=1}^{m} r_{\sigma(i)} \leq \oplus_{i=1}^{m} r_i = (x, x'), \quad \oplus_{i=1}^{m} r_{\sigma(i)} \sim (x, x'), \tag{5.3}$$

which is implied by the fact that $\mathcal{U}$ consists of intermediate species. Moreover, there exist $0 = k_0 < k_1 < \cdots < k_j < k_{j+1} = m$, such that for each $i = 0, \dots, j$, either $r_{\sigma(k_i+1)}, \dots, r_{\sigma(k_{i+1})} \in \mathcal{R}_0$ or $r_{\sigma(k_i+1)} \in \mathcal{R}'_{\mathcal{U}} \setminus \mathcal{R}_{\mathcal{U}}$, $r_{\sigma(k_i+1)}, \dots, r_{\sigma(k_{i+1}-1)} \in \mathcal{R}_{\mathcal{U}} \cap \mathcal{R}'_{\mathcal{U}}$ and $r_{\sigma(k_{i+1})} \in \mathcal{R}_{\mathcal{U}} \setminus \mathcal{R}'_{\mathcal{U}}$ with $r_{\sigma(k_i+1)} \oplus \dots \oplus r_{\sigma(k_{i+1})} \in \overline{\mathcal{R}_0}$. Therefore, $r_{\sigma(k_i+1)} \oplus \cdots \oplus r_{\sigma(k_{i+1})} \in \mathrm{cl}\big(\mathcal{R}^*_{\mathcal{U},\mathcal{F}}\big)$ for all $i = 1, \dots, j$. This yields $\oplus_{i=1}^{m} r_{\sigma(i)} \in \mathrm{cl}\big(\mathcal{R}^*_{\mathcal{U},\mathcal{F}}\big)$ as well. Combining (5.3) and Lemma 2, it follows that x leads to x' via $\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$. The proof is complete. □

Theorem 5.2(2) does not hold in general, not even for non-interacting species. Consider the following counterexample:

$$\mathcal{R} = \{S_1 \longrightarrow S_2 + U \longrightarrow S_3, S_2 \longrightarrow S_4, S_4 + U \longrightarrow S_5\}$$

with $\mathcal{U} = \{U\}$ and $\mathcal{F} = \mathcal{R}_{\mathcal{U}}$. Then,

$$\mathcal{R}^*_{\mathcal{U},\mathcal{F}} = \{S_2 \longrightarrow S_4, S_1 \longrightarrow S_3, S_1 + S_4 \longrightarrow S_2 + S_5\}.$$

Note that $(S_1, S_5) = (S_1 \longrightarrow S_2 + U) \oplus (S_2 \longrightarrow S_4) \oplus (S_4 + U \longrightarrow S_5)$. Thus, S_1 leads to S_5 via $\mathcal{R}$, but not via $\mathcal{R}^*_{\mathcal{U},\mathcal{F}}$.

6 Discussion and conclusion

We introduced and analysed the properties of a sum operation on chemical reactions. Thereby, we connect and characterise structural properties of RNs, such as reachability, (weakly)

https://doi.org/10.1017/S0956792522000146 Published online by Cambridge University Press

reversibility and being essential via the closure of the sum operation. This extends previous characterisations [7, 33, 39] and connects such properties to the geometry of the closure $\mathrm{cl}(\mathcal{R})$ in the product space $\mathbb{N}_0^n \times \mathbb{N}_0^n$. In another direction, we defined reductions of RNs by elimination of species from an RN by adding reactions. Those reductions originate from connections to the slow-fast limits of stochastic RNs [6]. Furthermore, we studied the conservation of (weakly) reversibility, when reachability of the original and the reduced network coincide in some sense.

As the discrete dynamics of Petri nets and vector addition systems correspond directly to dynamics of RNs [9], the developed theory pertains to those areas as well. Correspondingly, problems and questions from theoretical computer science relate to the notions we have introduced. As an example, an undecidable problem relating to Section 3 asks whether two RNs given by their reaction sets $\mathcal{R}_1, \mathcal{R}_2$ with initial values x_1, x_2, respectively, have the same reachability sets, that is, whether $\mathcal{R}_1(x_1) = \mathcal{R}_2(x_2)$ [20]. Another example is the decidable reachability problem that asks whether given an RN and two states x_1, x_2, we can reach x_2 from x_1 [9, 39].

Furthermore, the closure $\mathrm{cl}(\mathcal{R})$ of an RN has only sometimes the structure of a semi-linear set. This is not surprising as the set of reachable states of an RN directly relates to the closure $\mathrm{cl}(\mathcal{R})$ of $\mathcal{R}$, see Section 3. Reachability sets can be highly complex and are not necessarily semi-linear [22, 46]. Nonetheless, it might be interesting to characterise and study the structure of RNs $\mathcal{R}$ for which $\mathrm{cl}(\mathcal{R})$ is semi-linear.

Overall we hope that the sum calculus on reactions we have introduced will find further applications, possibly even in areas which a priori are not directly linked to our areas of research.

Acknowledgement

The work presented in this article is supported by Novo Nordisk Foundation, grant NNF19OC0058354. LH acknowledges funding from the Swiss National Science Foundation Early Postdoc.Mobility grant (P2FRP2_188023). Furthermore, we are grateful to the reviewers for helpful suggestions.

References

[1] Anderson, B. M. & May, R. M. (1991) *Infectious Diseases of Humans: Dynamics and Control*, 5th ed. Interdisciplinary Applied Mathematics, Vol. 17, Oxford University Press, Oxford.

[2] Anderson, D., Craciun, G. & Kurtz, T. (2010) Product-form stationary distributions for deficiency zero chemical reaction networks. *Bul. Math. Biol.* **72**, 1947–1970.

[3] Anderson, D. & Kurtz, T. (2011) Continuous time Markov Chain models for chemical reaction networks. In: H. Koeppl, D. Densmore, G. Setti and M. Di Bernardo (editors), *Design and Analysis of Biomolecular Circuits*, Springer-Verlag, New York, pp. 3–42.

[4] Anderson, D. F. & Cotter, S. L. (2016) Product-form stationary distributions for deficiency zero networks with non-mass action kinetics. *Bull. Math. Biol.* **78**, 2390–2407.

[5] Bowen, J., Acrivos, A. & Oppenheim, A. (1963) Singular perturbation refinement to quasi-steady state approximation in chemical kinetics. *Chem. Eng. Sci.* **18**(3), 177–188.

[6] Cappelletti, D. & Wiuf, C. (2016) Elimination of intermediate species in multiscale stochastic reaction networks. *Ann. Appl. Probab.* **26**(5), 2915–2958.

[7] Cappelletti, D. & Wiuf, C. (2016) Product-form Poisson-like distributions and complex balanced reaction systems. *SIAM J. Appl. Math.* **76**(1), 411–432.

[8] Christiansen, J. A. (1953) The elucidation of reaction mechanisms by the method of intermediates in quasi-stationary concentrations. *Adv. Catal.* **5**, 311–353.

[9] Cook, M., Soloveichik, D., Winfree, E. & Bruck, J. (2009) *Programmability of Chemical Reaction Networks*, Springer, Berlin, Heidelberg, pp. 543–584.
[10] Cornish-Bowden, A. (2013) *Fundamentals of Enzyme Kinetics*, Springer, New York.
[11] Ewens, W. (2004) *Mathematical Population Genetics 1: Theoretical Introduction*. Interdisciplinary Applied Mathematics, Springer, New York.
[12] Feinberg, M. (2019) *Foundations of Chemical Reaction Network Theory*, 1st ed. Interdisciplinary Applied Mathematics, Vol. 17, Springer, Cham.
[13] Feliu, E., Lax, C., Walcher, S. & Wiuf, C. (2019) Quasi-steady state and singular perturbation reduction for reaction networks with non-interacting species. arXiv preprint arXiv:1908.11270.
[14] Feliu, E. & Wiuf, C. (2012) Variable elimination in chemical reaction networks with mass-action kinetics. *SIAM J. Appl. Math.* **72**(4), 959–981.
[15] Feliu, E. & Wiuf, C. (2013) Simplifying biochemical models with intermediate species. *J. R. Soc. Interface* **10**, 87.
[16] Gardiner, C. W. (2004) *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 3rd ed. Springer Series in Synergetics, Vol. 13, Springer-Verlag, Berlin.
[17] Ginsburg, S. & Spanier, E. H. (1964) Bounded ALGOL-like languages. *Trans. Am. Math. Soc.* **113**(2), 333–368.
[18] Ginsburg, S. & Spanier, E. H. (1966) Semigroups, Presburger formulas, and languages. *Pacific J. Math.* **16**(2), 285–296.
[19] Gunawardena, J. (2012) A linear framework for time-scale separation in nonlinear biochemical systems. *PloS One* **7**(5), e36321.
[20] Hack, M. (1976) The equality problem for vector addition systems is undecidable. *Theor. Comput. Sci.* **2**(1), 77–95.
[21] Hoessly, L. & Wiuf, C. (2022) Fast reactions with non-interacting species in stochastic reaction networks. *Math. Biosci. Eng.* **19**, 2720–2749.
[22] Hopcroft, J. & Pansiot, J.-J. (1979) On the reachability problem for 5-dimensional vector addition systems. *Theor. Comput. Sci.* **8**(2), 135–159.
[23] Hornos, J., Schultz, D., Innocentini, G., Wang, J., Walczak, A., Onuchic, J. & Wolynes, P. (2005) Self-regulating gene: an exact solution. *Phys. Rev. E* **72**(5), 051907.
[24] Kang, H.-W. & Kurtz, T. G. (2013) Separation of time-scales and model reduction for stochastic reaction networks. *Ann. Appl. Probab.* **23**(2), 529–583.
[25] Kepler, T. & Elston, T. (2001) Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* **81**(6), 3116–3136.
[26] King, E. & Altman, C. (1956) A schematic method of deriving the rate laws for enzyme-catalyzed reactions. *J. Phys. Chem.* **60**, 1375–1378.
[27] Kurtz, T. G. (1981) *Approximation of Population Processes*. Society for Industrial and Applied Mathematics.
[28] Murray, J. D. (2002) *Mathematical Biology: I. An Introduction*, 3rd ed. Interdisciplinary Applied Mathematics, Vol. 17, Springer, New York.
[29] Pantea, C., Gupta, A., Rawlings, J. B. & Craciun, G. (2014) The QSSA in chemical kinetics: as taught and as practiced. In: *Discrete and Topological Models in Molecular Biology*, Springer, Berlin, Heidelberg, pp. 419–442.
[30] Parikh, R. J. (1961) Language–generating devices. Quarterly Progress Report, No. 60, Research Laboratory of Electronics, MIT, pp. 199–212.
[31] Parikh, R. J. (1966) On context-free languages. *J. ACM* **13**(4), 570–581.
[32] Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. (2015) Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979.
[33] Paulevé, L., Craciun, G. & Koeppl, H. (2014) Dynamical properties of discrete reaction networks. *J. Math. Biol.* **69**(1), 55–72.
[34] Peccoud, J. & Ycart, B. (1995) Markovian modeling of gene-product synthesis. *J. Peccoud, B. Ycart.* **48**(2), 222–234.
[35] Sáez, M., Feliu, E. & Wiuf, C. (2018) Graphical criteria for positive solutions to linear systems. *Linear Algebra Appl.* **552**, 166–193.

[36] SÁEZ, M., WIUF, C. & FELIU, E. (2017) Graphical reduction of reaction networks by linear elimination of species. *J. Math. Biol.* **74**(1), 195–237.
[37] SAKAMOTO, A. & KAWAKAMI, H. (1988) A graphical approach to complex reaction networks. *Chem. Phys. Lett.* **146**(5), 444–448.
[38] SINGHAL, G., RENGER, G., SOPORY, S. & IRRGANG, K. (2012) *Concepts in Photobiology: Photosynthesis and Photomorphogenesis*, Springer, Dordrecht.
[39] SZLOBODNYIK, G., SZEDERKÉNYI, G. & JOHNSTON, M. (2019) Reachability analysis of subconservative discrete chemical reaction networks. *Match* **81**, 705–736.
[40] TEMKIN, O. & BONCHEV, D. (1992) Application of graph theory to chemical kinetics: part 1. kinetics of complex reactions. *J. Chem. Edu.* **69**(7), 544–550.
[41] TEMKIN, O., ZEIGARNIK, A. W. & BONCHEV, D. G. (1996) *Chemical Reaction Networks: A Graph-Theoretical Approach*, CRC Press, Boca Raton.
[42] THATTAI, M. & VAN OUDENAARDEN, A. (2001) Intrinsic noise in gene regulatory networks. *PNAS* **98**(15), 8614–8619.
[43] WEIDLICH W. & GÜNTER H. (2012) *Concepts and Models of a Quantitative Sociology The Dynamics of Interacting Populations*. Springer Series in Synergetics, Springer, Berlin, Heidelberg.
[44] WILKINSON, D. (2006) *Stochastic Modelling for Systems Biology*, Chapman and Hall/CRC, Boca Raton.
[45] WIUF, C. & XU, C. (2020) Classification and threshold dynamics of stochastic reaction networks. arXiv preprint arXiv:2012.07954.
[46] YAMASAKI, H. (1984) Normal petri nets. *Theor. Comput. Sci.* **31**(3), 307–315.