computation

MDPI

Article

Simplification of Reaction Networks, Confluence and Elementary Modes †

Guillaume Madelaine [1], Elisa Tonello [2], Cédric Lhoussaine [1] and Joachim Niehren [1,3,*]

1 CRIStAL—Centre de Recherche en Informatique Signal et Automatique de Lille—CNRS UMR 9189, Université de Lille, F-59000 Lille, France; guillaume.madelaine@ed.univ-lille1.fr (G.M.); cedric.lhoussaine@univ-lille1.fr (C.L.)

2 School of Mathematical Sciences, University of Nottingham, NG7 2RD Nottingham, UK; Elisa.Tonello@nottingham.ac.uk

3 INRIA—French Institute for Research in Computer Science and Automation, 59000 Lille, France

* Correspondence: joachim.niehren@inria.fr

† This paper is an extended version of our paper published in the Proceedings of the International Conference on Computational Methods in Systems Biology 2016, Cambridge, UK, 21–23 September 2016.

Academic Editors: Gennady Bocharov, Olga Solovyova and Vitaly Volpert
Received: 17 January 2017; Accepted: 26 February 2017; Published: 10 March 2017

Abstract: Reaction networks can be simplified by eliminating linear intermediate species in partial steady states. In this paper, we study the question whether this rewrite procedure is confluent, so that for any given reaction network with kinetic constraints, a unique normal form will be obtained independently of the elimination order. We first show that confluence fails for the elimination of intermediates even without kinetics, if "dependent reactions" introduced by the simplification are not removed. This leads us to revising the simplification algorithm into a variant of the double description method for computing elementary modes, so that it keeps track of kinetic information. Folklore results on elementary modes imply the confluence of the revised simplification algorithm with respect to the network structure, i.e., the structure of fully simplified networks is unique. We show, however, that the kinetic rates assigned to the reactions may not be unique, and provide a biological example where two different simplified networks can be obtained. Finally, we give a criterion on the structure of the initial network that is sufficient to guarantee the confluence of both the structure and the kinetic rates.

Keywords: simplification; confluence; reaction network; ordinary differential equations; deterministic semantics; elementary modes; system biology; rewriting rules

1. Introduction

Chemical reaction networks are widely used in systems biology for modeling the dynamics of biochemical molecular systems [1–4]. A chemical reaction network has a graph structure that can be identified with an (unmarked) Petri net [5]. Beside of this, it assigns to each of its reactions a kinetic rate that models the reaction's speed. Chemical reaction networks can either be given a deterministic semantics in terms of ordinary differential equations (ODEs), which describes the evolution of the average concentrations of the species of the network over time, or a stochastic semantics in terms of continuous time Markov chains, which defines the evolution of molecule distributions of the different species over time. In this paper, we focus on the deterministic semantics.

Reaction networks modeling molecular biological systems—see, e.g., the examples in the BioModels database [6]—may become very large if modeling sufficient details. Therefore, biologists like to abstract whole subnetworks into single black-box reactions, usually in an adhoc manner that ignores kinetic information [7,8]. The absence or loss of kinetic information, however, limits the applicability

of formal analysis techniques. Therefore, much effort has been spent on simplification methods for reaction networks that preserve the kinetic information (see [9] for an overview).

The classical example for a structural simplification method is Michaelis-Menten's reduction of enzymatic networks with mass-action kinetics [10]. It removes the intermediate species—the complex C and enzyme E—under the assumption that their concentrations $C(t)$ and $E(t)$ are quasi steady, i.e., approximately constant for all time points t after a short initial phase. Segel [11] shows how to infer Michaelis-Menten's simplification from the assumptions that $C(t)$ is constant and that the conservation law $C(t) + E(t) = E(0)$ holds. This is equivalent to our exact steadiness assumption for both $C(t)$ and $E(t)$.

$$S + E \underset{k_2 C}{\overset{k_1 SE}{\rightleftharpoons}} C \xrightarrow{k_3 C} P + E \quad \text{simplifies to} \quad S \xrightarrow{k_3(E(0)+C(0))\frac{S}{\frac{k_2+k_3}{k_1}+S}} P$$

The ODEs for C inferred from this network jointly with exact steady state assumptions for C and E entail that the concentration of substrate S must be constant too, even if the network is used in a bigger context where the intermediate C is neither produced nor consumed. In the literature, this consequence is usually mentioned but ignored when considering the production rate of product P as a function of the concentration of S for the enzymatic network in isolation (see e.g., [12]). This oversimplification can be avoided when studying the enzymatic network in the context of a larger network. For instance, the steady state assumptions for C, E, and thus S can be satisfied in the context of the reaction network with the reaction $\emptyset \xrightarrow{k_4} S$ which produces S with constant speed k_4, and the reaction $P \xrightarrow{k_5 P} \emptyset$ which degrades P with mass-action kinetics with rate constant k_5. In this context, the concentration of P will saturate quickly under exact steady state assumptions for C, E, and thus S, as illustrated in Figure 1, while in other contexts it may grow without bound or even oscillate. The Michaelis-Menten simplification of the enzymatic network indeed preserves the dynamics of a network in any context which does not produce nor consume the intermediates E and C, under the assumption that E and C are exactly in steady state with respect to the network in the context.

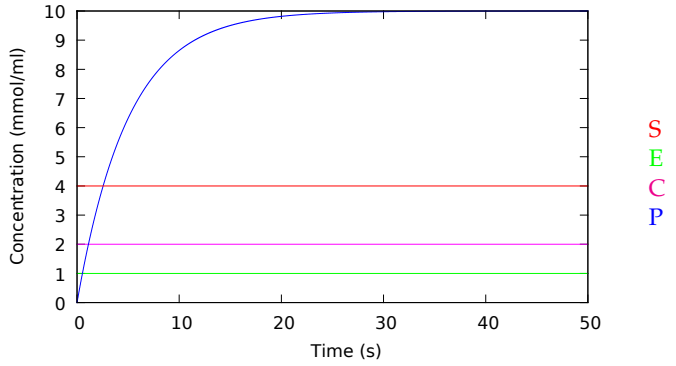


Figure 1. Evolution of the concentration of S, E, C and P in enzymatic network with mass-action kinetics with the parameters $k_1 = k_2 = k_3 = 1$, the initial concentrations $E(0) = 1$, $C(0) = 2$, $S(0) = 4$, $P(0) = 0$, and in the context of the network with a reaction $\emptyset \xrightarrow{k_4} S$ which produces S with constant speed $k_4 = 2$, and a reaction $P \xrightarrow{k_5 P} \emptyset$ which degrades P with parameter $k_5 = 0.2$.

Whether exact steady state assumptions are realistic is an interesting question since the concentrations may be at most close to steady in practice. In the literature it has been argued that the Michaelis-Menten simplification yields a good approximation under appropriate conditions [11,13,14], which typically depend on the context. Whether such properties can be extended to more general simplification methods as developed in the present article is an interesting question but out of the scope of the paper.

Alternatively, much work was spent on simplifying the ODEs inferred from a given reaction network [15,16], rather than the reaction network by itself. Indeed, any structural simplification method on the network level, that preserves the kinetic information with respect to the deterministic semantics, must induce a reduction method on the ODE level. The opposite must not be true, since some ODEs may not be derivable from any reaction network or may be inferred from many different ones [17]. Furthermore, it is not clear what it could mean for an ODE simplification method to be contextual. Therefore, ODE simplification alone cannot be understood as a simplification of biological systems.

A general structural simplification algorithm for reaction networks with deterministic semantics was first presented by Radulescu et al. They proposed yet another method [18] for simplifying reaction networks with kinetic expressions in partial steady states. Their method assumes the same linearity restriction considered in this paper, preserves exactly the deterministic semantics, but uses different algorithmic techniques. Their simplification algorithm is based on a graph of intermediate species. It computes cycles for simplifying the network structure rather than on elementary modes, and spanning trees for simplifying the kinetic expressions. A set of intermediate species is eliminated in one step, leading to a unique result, that is included in the results found with the algorithm of the present paper. We have not understood yet what distinguishes this result from the others obtained with our algorithm; a clarification of this point might shed light on the relationship between the two methods. In the same paper, the authors also observe that applying the method iteratively to intermediates one by one leads to different results even with different structure. The reason is that dependency elimination is lost in this manner.

A purely structural simplification algorithm method for reaction networks without kinetic rates was proposed in [19]. The method allows to remove some intermediate species by combining the reaction producing and consuming them. For instance, one can simplify the network with the following two reactions on the left into the single reaction on the right, by removing the intermediate species B:

$$\left.\begin{array}{l} A_1 + \ldots + A_n \to B \\ B \to C_1 + \ldots + C_m \end{array}\right\} \text{ simplifies to } A_1 + \ldots + A_n \to C_1 + \ldots + C_m$$

Since no partial steady state assumptions can be imposed in a kinetics free framework, the intermediate elimination rules need some further restrictions. Given these, the simplification steps were shown correct with respect to the attractor semantics. contextual equivalence relation was obtained by instantiating the general framework for observational program semantics from [20]. Rather than being based on termination as observable for concurrent programs, it relies on the asymptotic behaviours of the networks represented by the terminal connected components, which are often called attractors.

Outline

We first recall some basic notions on confluence, multisets, and commutative semigroups in Section 2. In Section 3 we recall the basics on reaction networks without kinetics and elementary flux modes. In Section 4, we present the rewrite rules for intermediate elimination, illustrate the failure of confluence, and propose a rewrite rule for eliminating dependent reactions, which however turns out to be non-confluent on its own. In Section 5 we present the refined algorithm in the case without kinetics based on the notion of flux networks for representing reaction networks, and prove its confluence by reduction to a folklore result on elementary flux modes. In Section 6, we introduce reaction networks with kinetic expressions, and extend them with kinetic constraints. In Section 7, we lift the revised algorithm to constrained flux networks with kinetics. In Section 8 we present a linearity restriction, that is preserved by reductions, and thus structurally confluent. In Section 9, we present a counter example that shows that full confluence is still not achieved, and present a further syntactic restriction based on elementary modes avoiding this problem. Section 10 provides a biological example of non-confluence with kinetics. Section 11 studies the relation between the simplification and the underlying ODEs simplification. Finally, we conclude in Section 12.

2. Preliminaries

We recall basic notions on confluence of binary relations, on multisets, and more general commutative semigroups. We will denote the set of all natural numbers including 0 by $\mathbb{N}$ and the set of integers by $\mathbb{Z}$.

2.1. Confluence Notions

We recall the main confluence notions and their relationships from the literature.

Let $(S, \sim)$ be a set with an equivalence relation and $\to \subseteq S \times S$ a binary relation. In most cases, $\sim$ will be chosen as the equality relation of the set S, which is $=^S = \{(s, s) \mid s \in S\}$. We define $\to^0 = \sim$ and $\to^k = \to \circ \to^{k-1}$ for all $k \in \mathbb{N} \setminus \{0\}$. The relation $\to^* = \bigcup_{k \in \mathbb{N}} \to^k$ is called the reflexive transitive closure of $\to$.

Definition 1. *We say that a binary relation $\to$ on $(S, \sim)$ is* confluent *if $\leftarrow^* \circ \to^* \subseteq \to^* \circ {}^*\leftarrow$ and* locally confluent *if $\leftarrow \circ \to \subseteq \to^* \circ {}^*\leftarrow$. We say that two binary relations $\Rightarrow$ and $\to$ on S commute if $\Leftarrow \circ \to \subseteq \to \circ \Leftarrow$.*

The confluence notions are illustrated by the diagrams in Figure 2. Clearly, a confluence of relation $\to$ is confluent if its reflexive transitive closure $\to^*$ commutes with itself. It is also obvious that local confluence implies confluence, and well known that the converse does not hold. In this paper, we will always use binary relations that are terminating, i.e., for any $s \in S$ there exists a $k \in \mathbb{N}$ such that $\{s' \mid s \to^k s'\} = \emptyset$, i.e., the length k of sequences of reduction steps starting with s is bounded. It is well known that locally confluent and terminating relations are confluent (Newman's lemma).

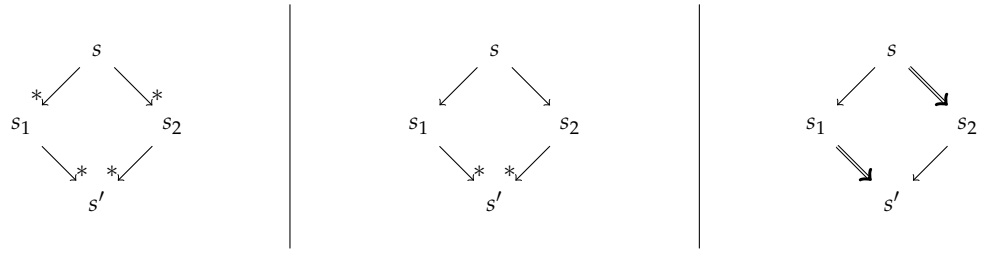


Figure 2. Confluence, local confluence, and commutation.

Lemma 1. *If a binary relation $\to$ on $(S, \sim)$ is confluent and commutes with $\sim$, then the binary relation $\sim \circ \to \circ \sim$ on $(S, =_S)$ is confluent.*

Definition 2. *Let $(S, \sim, \to)$ and $(S', \approx, \Rightarrow)$ be two sets each endowed with two binary relations. A function $T : S \to S'$ is called a* simulation *from $(S, \sim, \to)$ to $(S', \approx, \Rightarrow)$ if for any $s_1, s_2 \in S$, if $s_1 \sim s_2$ then $T(s_1) \approx T(s_2)$, and if $s_1 \to s_2$ then $T(s_1) \Rightarrow T(s_2)$.*

The conditions that have to be satisfied by simulations are illustrated by the diagrams in Figure 3.

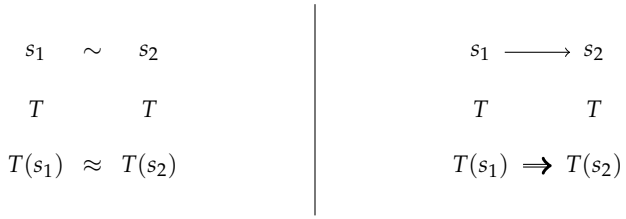


Figure 3. Simulation diagrams.

2.2. Multisets

Let R be a finite set. A multiset M with elements in R is a function $M : R \to \mathbb{N}$. For any $r \in R$ we call $M(r)$ the number of occurrences of r in M. We say that r is a member of multiset M and write $r \in M$ if $M(r) \neq 0$. We denote by $\mathcal{M}_R$ the set of all multisets (over R), and will simply write $\mathcal{M}$ if the set R is clear from the context.

Given numbers $k, n_1, \ldots, n_k \in \mathbb{N}$ and a subset $\{r_1, \ldots, r_k\} \subseteq R$ with k different elements, we denote by $M = n_1 r_1 + \cdots + n_k r_k = \sum_{i=1}^{n} n_i r_i$ the multiset that for any $1 \leq i \leq k$ contains $M(r_i) = n_i$ occurrences of r_i and $M(r) = 0$ occurrences of all other elements in R.

The sum of two multisets $M_1 +^{\mathcal{M}} M_2$ is the multiset M that satisfies $M(r) = M_1(r) +^{\mathbb{N}} M_2(r)$ for all $r \in R$. The empty multiset $0^{\mathcal{M}}$ is the function that maps all elements of R to 0. The algebra of multisets $(\mathcal{M}, +^{\mathcal{M}}, 0^{\mathcal{M}})$ over a given set R is a commutative semigroup with a neutral element.

It should be noticed that our notation may give rise to some ambiguities, since we will also write $+$ for the addition of natural numbers instead of $+^{\mathbb{N}}$. This may be problematics if $R = \mathbb{N}$. In this case, the notation introduced below we will permit us to write $(n_1 r_1 + \cdots + n_k r_k)^{\mathcal{M}} = (\sum_{i=1}^{n} n_i r_i)^{\mathcal{M}}$ for sums of multisets and $(n_1 r_1 + \cdots + n_k r_k)^{\mathbb{N}} = (\sum_{i=1}^{n} n_i r_i)^{\mathbb{N}}$ for sums of natural numbers.

2.3. Commutative Semigroups

Let $(\mathcal{G}, +^{\mathcal{G}}, 0^{\mathcal{G}})$ and $(\mathcal{F}, +^{\mathcal{F}}, 0^{\mathcal{F}})$ be two semigroups with neutral element. Beside of the algebras of multisets (depending on the choice of R) we are interested in the algebra of vectors of naturals $(\mathbb{N}^n, +^{\mathbb{N}^n}, 0^{\mathbb{N}^n})$ for any $n \in \mathbb{N}$.

A homomorphism between two semigroups is a function $h : \mathcal{G} \to \mathcal{F}$ such that $h(g_1 +^{\mathcal{G}} g_2) = h(g_1) +^{\mathcal{F}} h(g_2)$ for all $g_1, g_2 \in \mathcal{G}$ and $h(0^{\mathcal{G}}) = 0^{\mathcal{F}}$. A homomorphism $h : \mathcal{M}_R \to \mathcal{F}$ on multisets is determined by the values of h on singleton multisets in $\mathcal{M}_R$ via the equation:

$$h(n_1 r_1 + \ldots + n_k r_k) = \underbrace{h(1r_1) +^{\mathcal{F}} \ldots +^{\mathcal{F}} h(1r_1)}_{n_1 \text{ times}} +^{\mathcal{F}} \ldots +^{\mathcal{F}} \underbrace{h(1r_k) +^{\mathcal{F}} \ldots +^{\mathcal{F}} h(1r_k)}_{n_k \text{ times}}.$$

Given a homomorphism $h : \mathcal{M}_R \to \mathcal{F}$, we define the interpretation $M^{\mathcal{F}} = h(M)$ for all multisets $M \in \mathcal{M}_R$. Clearly, the interpretation depends on the homomorphism h, even though only its co-domain $\mathcal{F}$ appears in our notation. This works smoothly since there will never be any ambiguity about the homomophism that is chosen. If $R = \mathcal{F}$, then we use the homomorphism $eval_{\mathcal{F}} : \mathcal{M}_{\mathcal{F}} \to \mathcal{F}$ with $eval_{\mathcal{F}}(1f) = f$ for all elements $f \in \mathcal{F}$. In this case, any multiset $n_1 f_1 + \ldots + n_k f_k$ with elements in $\mathcal{F}$ is evaluated to a single element $(n_1 f_1 + \ldots + n_k f_k)^{\mathcal{F}} = eval_{\mathcal{F}}(n_1 f_1 + \ldots + n_k f_k)$ and thus by the above equation:

$$(n_1 f_1 + \ldots + n_k f_k)^{\mathcal{F}} = \underbrace{f_1 +^{\mathcal{F}} \ldots +^{\mathcal{F}} f_1}_{n_1 \text{ times}} +^{\mathcal{F}} \ldots +^{\mathcal{F}} \underbrace{f_k +^{\mathcal{F}} \ldots +^{\mathcal{F}} f_k}_{n_k \text{ times}}.$$

If $\mathcal{F} = \mathcal{M}_R$ then we use the identity homomorphism $id_{\mathcal{M}_R} : \mathcal{M}_R \to \mathcal{M}_R$ with $id_{\mathcal{M}_R}(M) = M$ for all $M \in \mathcal{M}_R$. In this case we have that $(n_1 r_1 + \ldots + n_k r_k)^{\mathcal{M}_R} = n_1 r_1 + \ldots + n_k r_k$ is the multiset itself.

We will also use this notation in order to distinguish the operator $+$ of multisets in $\mathcal{M}_{\mathbb{N}}$ from the operator $+$ of natural numbers in $\mathbb{N}$ which we overloaded (as stated earlier). For instance, if $n, m \in \mathbb{N}$, then $(2n + 5m)^{\mathcal{M}_{\mathbb{N}}}$ is a multiset of natural numbers while $(2n + 5m)^{\mathbb{N}}$ is a natural number. Note also that different multisets may have the same interpretation. For instance if $n = 3$ and $m = 4$, then $(2n + 5m)^{\mathbb{N}} = 26 = (n2 + m5)^{\mathbb{N}}$ where we use $eval_{\mathbb{N}}$ as homomorphism while $(2n + 5m)^{\mathcal{M}_{\mathbb{N}}} \neq (n2 + m5)^{\mathcal{M}_{\mathbb{N}}}$ where we use $id_{\mathcal{M}_{\mathbb{N}}}$ as homomorphism.

For any subset $G \subseteq \mathcal{G}$ of a semigroup, we can define the (positive integer convex) cone of G, as the set of all positive integer linear combinations of elements of G:

$$cone(G) = \{(n_1 g_1 + ... + n_k g_k)^{\mathcal{G}} \mid k \in \mathbb{N},\ g_1 ... g_k \in G,\ n_1 ... n_k \in \mathbb{N}\}.$$

Here we use $eval_{\mathcal{G}}$ as homomorphism.

3. Reaction Networks without Kinetics

Let *Spec* be a finite set of *species* that is totally ordered. A *(chemical) solution* with species in *Spec* is a multiset of species $s : Spec \to \mathbb{N}$. A *(chemical) reaction* with species in *Spec* is a function $r : Spec \to \mathbb{Z}$, which assigns to each species A the stoichiometry of A in r. A chemical reaction r consumes the chemical solution $Cons_r = -r_{|\{A \in Spec | r(A) < 0\}}$ and produces the chemical solution $Prod_r = r_{|\{A \in Spec | r(A) > 0\}}$. Clearly $r(A) = Prod_r(A) - Cons_r(A)$ for all species A, while $Cons_r$ and $Prod_r$ are disjoint multisets in chemical reactions r (since their definition is based on stoichiometries).

We will freely identify a reaction r with the pair of chemical solutions consumed and produced by r. We will denote such pairs as $Cons_r \to Prod_r$. For instance, $B + 2C \to A$ is the chemical reaction r with $r(A) = 1$, $r(B) = -1$, and $r(C) = -2$. Note also that we do not consider $2A + B \to 3A + 2C$ as a chemical reaction, since the species A belongs to the chemical solutions on both sides. When removing $2A$ on both sides, we obtain a chemical reaction $B \to A + 2C$. The rewrite relation of a chemical reaction r contains all pairs of chemical solutions (s, s') such that $s'(A) = s(A) +^{\mathbb{N}} r(A)$ for all species A.

Definition 3. *A* reaction network (without kinetics) *over Spec is a finite set of chemical reactions over Spec, with a total order.*

To any reaction network N with total order $<$ we assign a unique vector of reactions $\mathbf{r} = (r_1, \ldots, r_n)$ such that $N = \{r_1, \ldots, r_n\}$ and $r_1 < \ldots < r_n$. Conversely, for any tuple of distinct reactions $\mathbf{r} = (r_1, \ldots, r_n)$, we write $\mathtt{N}_{\mathbf{r}}$ for the reaction network $\{r_1, \ldots, r_n\}$ with the total order $r_1 < \ldots < r_n$.

Any reaction network can be represented by a bipartite graph as for a a Petri net, with a node for each species and a node of a different type for each reaction. We will draw species nodes with ovals and reaction nodes with squares. An arrow labeled by k from the node of a species A to the node of a reaction r means that A is consumed k times by r, i.e., $r(A) = -k$. Conversely, an arrow with label k from the node of a reaction r to the node of a species A means that A is produced k times by r, i.e., $r(A) = k$. We will freely omit the labels $k = 1$.

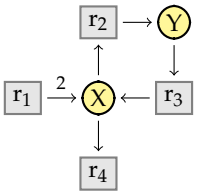
Example 1. *Consider the reaction network presented in Figure 4. It has $m = 2$ species $Spec = \{X, Y\}$ and $n = 4$ reactions $\{r_1, \ldots, r_4\}$ in that order. Reaction r_1 produces two molecules of species X out of nothing, reaction r_2 transforms an X into a molecule Y, while r_3 transforms a molecule Y back into a molecule X. Reaction r_4 degrades a molecule X.*

The set of chemical reactions defines an algebra $(\mathcal{R}, +^{\mathcal{R}}, 0^{\mathcal{R}})$ where $0^{\mathcal{R}}$ is the empty reaction $\to$, and $+^{\mathcal{R}}$ is the addition of integer valued functions on *Spec*. Note that $s' \to s +^{\mathcal{R}} s \to s' = 0^{\mathcal{R}}$ for any two disjoint chemical solutions s and s'. By interpretation in this algebra (that is using the identity homomorphism), we can evaluate each multiset of chemical reactions M as a chemical reaction $M^{\mathcal{R}}$ itself, as shown in Section 2.2.

Definition 4. *An* invariant *of a reaction network N without kinetics is a multiset M of reactions of N such that $M^{\mathcal{R}} = 0^{\mathcal{R}}$. We denote the set of all invariants of N by $inv(N)$.*

The reaction network in Figure 4 has the set of invariants $\{(n_1 M_1 + n_2 M_2)^{\mathcal{M}} \mid n_1, n_2 \in \mathbb{N}\}$ where $M_1 = r_1 + 2r_4$ and $M_2 = r_2 + r_3$. We next relate the notion of invariants of a reaction network to the kernel of its stoichiometry matrix.

$r_1 = \quad \rightarrow 2X$
$r_2 = X \rightarrow Y$
$r_3 = Y \rightarrow X$
$r_4 = X \rightarrow$



	r_1	r_2	r_3	r_4
X	2	−1	1	−1
Y	0	1	−1	0

Figure 4. A reaction network and the associated graph and stoichiometry matrix.

3.1. Stoichiometry Matrices

The stoichiometry information of a reaction network is usually collected in its stoichiometry matrix. For this we consider a set of species $Spec = \{A_1, \ldots, A_m\}$ and a reaction network $N = \{r_1, \ldots, r_n\}$, such that both sets are totally ordered by the indices of their elements.

The stoichiometry matrix S of N is the $m \times n$ matrix of integers, such that the entry of S at row i and column j is equal to $r_j(A_i)$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$. Note that reaction r_j contributes in the $j'th$ column, while species A_j contributes the j's row of S. For instance, the stoichiometry matrix of the reaction network in Figure 4 is given on the right.

It can now be noticed that, for any vector $v = (n_1, \ldots, n_n)$ of natural numbers, the multiset $n_1 r_1 + ... + n_n r_n$ is an invariant of reaction network N if and only if its stoichiometry matrix satisfies $Sv = 0$, i.e., if v belongs to the kernel of the stoichiometry matrix. Therefore, we define the (positive integer) kernel of a matrix S by:

$$ker_+(S) = \{v \in \mathbb{N}^n \mid Sv = 0\}.$$

3.2. Elementary Modes

The support of a vector $v = (n_1, \ldots, n_k)$ is the subset of indices i such that n_i is non-null, i.e., $supp(v) = \{i \in \{1, \ldots, k\} \mid n_i \neq 0\}$.

Definition 5. *An* elementary mode *of an $m \times n$ matrix S over $\mathbb{Z}$ is a vector $v \in ker_+(S) \setminus \{0^{\mathbb{N}^n}\}$ such that:*

v is on an extreme ray: *there exists no $v' \in ker_+(S) \setminus \{0^{\mathbb{N}^n}\}$ such that $supp(v') \subsetneq supp(v)$, and*
v is factorised: *there exists no $v'' \in ker_+(S)$ such that $v = kv''$ for some natural number $k \geq 2$.*

The condition $v \in ker_+(S)$ means that an elementary mode must be a (positive integer) steady state of S. Geometrically, the set of all positive integer steady states forms a pointed cone, that is generated by convex combinations of its extreme rays. The first condition states that any elementary flux mode v must belong to some extreme ray of the cone. The second condition requires that an elementary mode is maximally factorised, i.e., it is the vector on the extreme ray with the smallest norm.

Theorem 1 (Folklore [21])**.** *Let S be an $m \times n$ matrix of integers. Then the set E of all elementary modes of S has finite cardinality and satisfies $ker_+(S) = cone(E)$.*

The intuition is $ker_+(S)$ is a cone with a finite number of extreme rays, so that these extreme ray generated the cone. The set of elementary modes E contains exactly one point on each of the extreme rays of $ker_+(S)$. Therefore, $ker_+(S) = cone(E)$, i.e., the set of elementary modes is a finite generator of $ker_+(S)$.

Let us point out two differences between the definition of elementary mode considered here and in [21]. First, we added condition 2. Without this condition, any multiple of an elementary mode would be an elementary mode, so that there would be infinitely many. The double-description method as recalled there, however, computes the set of elementary modes in the above sense, so this difference is

minor. Second, note that [21] considers a slightly more general problem, where some of the coordinates of v may be negative. This corresponds to the addition of reversible reactions that we do not consider in the present paper.

3.3. Elementary Flux Modes

We next lift the concept of elementary modes from matrices to reaction networks, via the stoichiometry matrix. Given a vector of reactions $\mathbf{r} = (r_1, \ldots, r_n)$ and a vector $v = (n_1, \ldots, n_n)$ of natural numbers we define the multiset of reactions $v\mathbf{r}$ and the corresponding reaction $\mathbf{r}_v$ as follows:

$$v\mathbf{r} = n_1 r_1 + \ldots + n_n r_n \qquad \text{and} \qquad \mathbf{r}_v = (v\mathbf{r})^{\mathcal{R}}.$$

Definition 6. *An elementary flux mode of a reaction network $N = \mathbb{N}_\mathbf{r}$ is a multiset of reactions $v\mathbf{r}$ such that the vector v is an elementary mode of the stoichiometry matrix of N.*

The kernel condition $v \in ker_+(S)$ of elementary modes v yields that any elementary flux mode $v\mathbf{r}$ satisfies $\mathbf{r}_v = 0^{\mathcal{R}}$, i.e., the reaction defined by the elementary flux mode must be empty. For instance, reconsider the reaction network in Example 1 with $m = 2$ species and $n = 4$ reactions $\mathbf{r} = (r_1, r_2, r_3, r_4)$ in that order. Its stoichiometry matrix has two elementary modes: the vectors $v_1 = (1, 0, 0, 2)$ and $v_2 = (0, 1, 1, 0)$. The corresponding elementary flux modes are the multisets of reactions $v_1\mathbf{r} = r_1 + 2r_4$ and $v_2\mathbf{r} = r_2 + r_3$ illustrated in Figure 5 by the arrows coloured in apricot and aquamarine respectively. First consider the multiset $r_1 + 2r_4$: the first reaction r_1 produces $2X$ which are then degraded by $2r_4$. So the reaction $\mathbf{r}_{v_1} = (r_1 + 2r_4)^{\mathcal{R}} = 0^{\mathcal{R}}$ is indeed empty. Consider now the multiset of reactions $r_2 + r_3$: its first reaction r_2 transforms X to Y and its second reaction r_3 does the inverse. Thus, $\mathbf{r}_{v_2} = (r_2 + r_3)^{\mathcal{R}} = 0^{\mathcal{R}}$ is the empty reaction too. The intuition is that applying to a chemical solution at the same time all reactions of an elementary flux mode with their multiplicities does not have any effect.

$$\begin{array}{rcl@{\qquad}rcl} v_1 & = & (1,0,0,2) & v_1\mathbf{r} & = & r_1 + 2r_4 \\ v_2 & = & (0,1,1,0) & v_2\mathbf{r} & = & r_2 + r_3 \end{array}$$

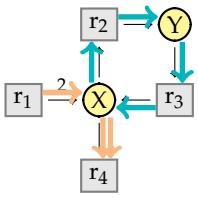


Figure 5. The elementary modes of the reaction network in Figure 4.

It should be noticed that the vector $v = (1, 1, 1, 2)$ is also a solution of the steady state equation $Sv = 0$, and thus the multiset of reactions $r_1 + r_2 + r_3 + 2r_4$ is also an invariant of the example network. It is the multiset sum of two elementary flux modes $v_1\mathbf{r} +^{\mathcal{M}} v_2\mathbf{r}$ which is also equal to $(v_1 +^{\mathbb{N}^4} v_2)\mathbf{r}$.

4. Simplifying Reaction Networks without Kinetics

We study the question whether the step-by-step intermediate elimination relation proposed in [22] is confluent in the case of reaction networks without kinetics. We present a counter example against the confluence and illustrate the reason for this problem.

4.1. Intermediate Elimination

Let $\mathcal{I} \subseteq Spec$ be a finite set of species that we will call *intermediate species* or *intermediates* for short. The simplification procedure will remove all intermediates from a given reaction network, step-by-step and in arbitrary order.

Our objective is to remove an intermediate $X \in \mathcal{I}$ from a network N by merging any pair of reactions of N, a reaction r that produces X and another reaction r' that consumes it. This is done

by the (INTER) rule in Figure 6, and is based on the merge operation $r \diamond_X r'$ which returns a linear combination of r and r' and thus of the reactions in the initial network:

$$r \diamond_X r' = (-r'(X)r + r(X)r')^{\mathcal{R}}.$$

Since r produces $r(X)$ molecules X while r' consumes $r'(X)$ molecules X, we have $(r \diamond_X r')(X) = 0$. Therefore, X is not present in the solutions consumed and produced by reaction $r \diamond_X r'$.

$$\text{(INTER)} \quad \frac{X \in \mathcal{I}}{N \Rrightarrow_{\text{Inter}} \cup\{r \in N \mid r(X) = 0\}}$$

$$\text{(DEP)} \quad \frac{k \in \mathbb{N} \qquad \mathbf{s} \in N^k \qquad v \in \mathbb{N}^k}{N \uplus \{\mathbf{s}_v\} \Rrightarrow_{\text{DEP}} N}$$

Figure 6. Simplification of reaction networks without kinetics with respect to a set $\mathcal{I}$ of intermediate species.

In Example 2 below, we will denote vectors $(n_1, \ldots, n_n)$ of natural numbers by $1^{n_1} \ldots n^{n_n}$, while freely omitting components i^{n_i} with $n_i = 0$ and simplifying component j^1 to j. For instance if $\mathbf{r} = (r_1, \ldots, r_4)$, we can write $\mathbf{r}_{14^2}$ instead of $\mathbf{r}_{(1,0,0,2)}$.

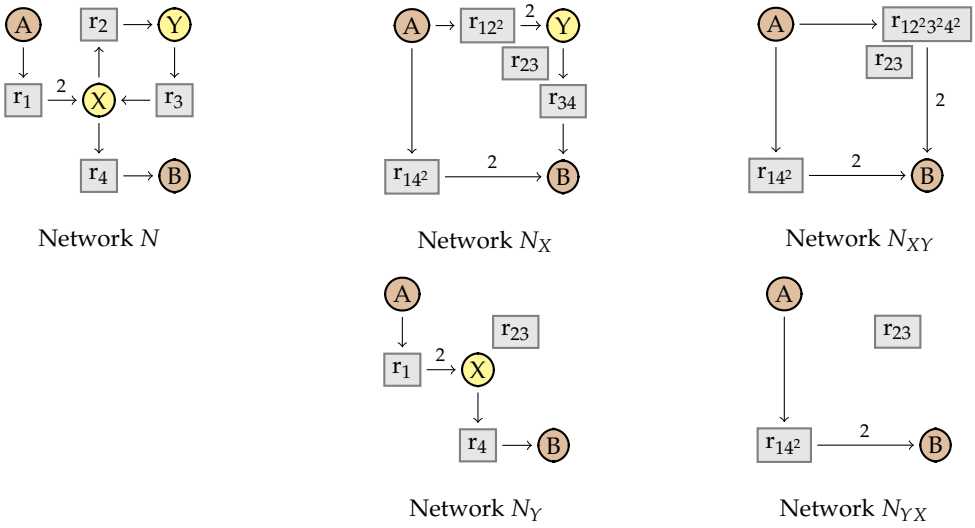


Figure 7. Elimination of intermediates X and Y in reaction network N in both possible orders, leading to two different final results N_{XY} and N_{YX}.

Example 2. *We consider the network N in Figure 7 with species $Spec = \{A, B, X, Y\}$ and reaction vector $\boldsymbol{r} = (r_1, \ldots, r_4)$. We consider the elimination of the intermediates in $\mathcal{I} = \{X, Y\}$ in both possible orders. On the top, we first eliminate the intermediate species X from N, obtaining network N_X. We have to combine reaction r_1 producing 2 X molecules with reaction r_2 which consumes 1 X molecule. We obtain the reaction $r_1 \diamond_X r_2 = \boldsymbol{r}_{12^2} = (r_1 + 2r_2)^{\mathcal{R}}$ that transforms one A molecule into 2 Y molecules. We proceed in the same way with the other 3 pairs of reactions that produce and consume X. Then, we can remove the intermediate species Y from network N_X and obtain the network N_{XY} in the top right. Note that we keep empty reactions such as $\boldsymbol{r}_{23}$. At the bottom, we show network N_Y, obtained by eliminating the intermediate species Y first. The only reaction producing Y in N is r_2 and the only reaction consuming Y is r_3. Merging them produces reaction $\boldsymbol{r}_{23}$. When eliminating intermediate X from N_Y, we obtain network N_{YX} on the bottom right.*

It turns out that N_{XY} and N_{YX} differ in that the former contains the reaction $\boldsymbol{r}_{12^23^24^2}$ in addition to the reactions $\boldsymbol{r}_{14^2}$ and $\boldsymbol{r}_{23}$ shared by both networks.

4.2. Eliminating Dependent Reactions

Example 2 shows that intermediate elimination with the (INTER) rule alone is not confluent, given that it may produce two different networks that cannot be simplified any further, N_{XY} and N_{YX}, depending on whether we first eliminate the intermediate X or the intermediate Y. The reaction network N_{XY} contains an additional reaction, which is a linear combination of two other reactions:

$$\mathbf{r}_{12^23^24^2} = (\mathbf{r}_{14^2} + 2\mathbf{r}_{23})^{\mathcal{R}}.$$

In order to solve this non-confluence problem, we propose the new simplification rule (DEP) in Figure 6. It eliminates a reaction that is a positive linear combination of other reactions of the network, i.e., some reaction $\mathbf{s}_v = (n_1 r_1 + \ldots + n_k r_k)^{\mathcal{R}}$ where $\mathbf{s} = (r_1, \ldots, r_k) \in N^k$ and $v = (n_1, \ldots, n_k) \in \mathbb{N}^k$ for some $k \in \mathbb{N}$.

Unfortunately, the simplification relation with rules (INTER) and (DEP) is still not confluent. The problem is that even applying rule (DEP) alone fails to be confluent as shown by the following counter example.

Example 3. *Consider the network N'' in Figure 8 in the absence of intermediates, i.e., where $\mathcal{I} = \emptyset$. There are two ways of applying rule* (DEP) *to this network, since $r_4 = (r_1 + 2r_2)^{\mathcal{R}}$ and $r_2 = (r_3 + r_4)^{\mathcal{R}}$. We can thus either eliminate r_4 leading to N''_{r_4} or r_2 leading to N''_{r_2}. The two results are different even though they contain no more dependencies.*

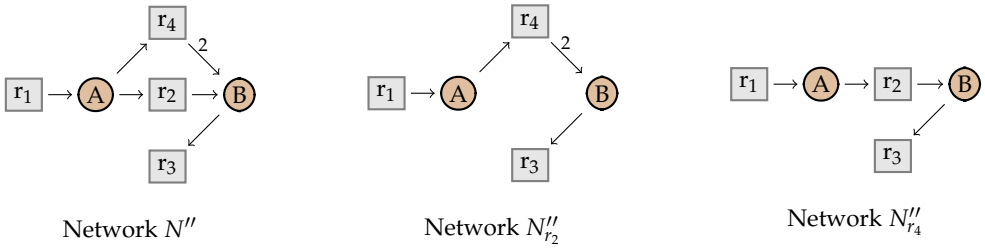


Figure 8. Dependency elimination is not confluent.

This example shows that general dependency elimination cannot be done in a confluent manner. On the other hand, what we need in order to solve the confluence problem for intermediate elimination as illustrated in Figure 7, is a little more restricted: it is sufficient to remove those dependent reactions that were introduced by intermediate elimination. Such dependencies can be identified from the vectors of natural numbers that we used to name the reactions. In the example, we have $\mathbf{r}_{12^23^24^2} = (\mathbf{r}_{14^2} + 2\mathbf{r}_{23})^{\mathcal{R}}$, so the dependency of this reaction follows from the dependency of the vectors $12^23^24^2 = ((14^2) + 2(23))^{\mathbb{N}^4}$.

5. Simplifying Flux Networks

We next introduce vector representations of reaction networks without kinetics, called flux networks, and show that the simplification of such representations can indeed be done in a confluent manner.

For the reminder of this section, we fix an n-tuple $\mathbf{r}$ of distinct reactions and a subset of species $\mathcal{I} \subseteq Spec$.

5.1. Vector Representations of Reaction Networks

The objective is to simplify the initial reaction network $\mathtt{N}_{\mathbf{r}}$ by removing the intermediates from $\mathcal{I}$. The iterative elimination of intermediate species generates a sequence of networks with reactions in

$cone(\mathtt{N}_{\mathbf{r}}) = \{\mathbf{r}_v \mid v \in \mathbb{N}^n\}$. The idea is now to use the vectors $v \in \mathbb{N}^n$ as representations of reactions $\mathbf{r}_v$. These vectors will tell us about the provenance of the reaction obtained when simplifying the network $\mathtt{N}_{\mathbf{r}}$.

The mapping of vectors $v \in \mathbb{N}^n$ to reactions $\mathbf{r}_v \in \mathcal{R}$ is a homomorphism between commutative semigroups, whose image is $cone(\mathtt{N}_{\mathbf{r}})$. It should be noticed, however, that it is not an isomorphism since any element of $ker_+(S)$ will be mapped to $0^{\mathcal{R}}$, where S is the stoichiometry matrix of $\mathtt{N}_{\mathbf{r}}$. Therefore, it makes a difference whether we will work with vectors in $\mathbb{N}^n$ representing a reaction or with the reactions itself. Intuitively, the difference is that we know where the reaction does come from.

Definition 7. *An n-ary flux network V is a finite subset of vectors in $\mathbb{N}^n$ that is totally ordered.*

Any n-ary flux network V defines a reaction network $\mathbf{r}_V = \{\mathbf{r}_v \mid v \in V\}$, that we call the reaction network represented by V. The total order of the reactions in network $\mathbf{r}_V$ is the one induced by the total order of V.

5.2. Simplification Rules

Let $\mathcal{I} \subseteq Spec$ be a finite set of species that we call intermediates. In Figure 9, we rewrite the simplification rules (F-INTER) and (F-DEP) so that they apply to flux networks. For this we have to lift the merge operation from reactions to vectors that represent them. For any $v_1, v_2 \in \mathbb{N}^n$ we define:

$$v \diamond_X v' = (-\mathbf{r}_{v'}(X)v + \mathbf{r}_v(X)v')^{\mathbb{N}^n}.$$

In the rule for the dependency elimination, we now use a notation for linear combinations of vectors in $\mathbb{N}^n$. Given a vector $\mathbf{v} = (v_1, \ldots, v_k)$ of vectors in $\mathbb{N}^n$ and a vector $v = (n_1, \ldots, n_k)$ of natural numbers we define:

$$\mathbf{v}_v = (n_1 v_1 + \ldots + n_k v_k)^{\mathbb{N}^n}.$$

The counter example for the non-confluence of dependency elimination can no more be applied in this way, since rule (F-DEP) is not based on the dependency of the reactions as with (DEP) but on the dependencies of the vectors that define the reactions.

$$\text{(F-INTER)} \quad \frac{X \in \mathcal{I} \qquad \{v \in V \mid \mathbf{r}_v(X) \neq 0\} \neq \emptyset}{V \Rrightarrow_{\text{F-INTER}} \begin{array}{l}\{v \diamond_X v' \mid v, v' \in V,\ \mathbf{r}_v(X) > 0,\ \mathbf{r}_{v'}(X) < 0\} \\ \cup \{v \in V \mid \mathbf{r}_v(X) = 0\}\end{array}}$$

$$\text{(F-DEP)} \quad \frac{k \in \mathbb{N} \qquad \mathbf{v} \in V^k \qquad v \in \mathbb{N}^k}{V \uplus \{\mathbf{v}_v\} \Rrightarrow_{\text{F-DEP}} V}$$

$$\text{(F-FACT)} \quad \frac{v \in \mathbb{N}^n \setminus \{0^{\mathbb{N}^n}\} \qquad k \geq 2}{V \uplus \{(kv)^{\mathbb{N}^n}\} \Rrightarrow_{\text{F-FACT}} V \cup \{v\}}$$

Figure 9. Simplifying flux networks for an initial n-tuple of reactions $\mathbf{r}$ and a set of intermediate species $\mathcal{I}$.

5.3. Factorization

The simplification relation with axioms (F-INTER) and (F-DEP) is still not confluent, as shown in Example 4.

Example 4. *We consider the vector of initial reactions $\boldsymbol{r} = (r_1, r_2, r_3)$ of the network N'' in Figure 10. Let $\mathcal{I} = \{X, Y, Z\}$ be the set of intermediate species. Note that $N'' = \boldsymbol{r}_{V_3}$ where $V_3 = \{(1,0,0), (0,1,0), (0,0,1)\} \subseteq \mathbb{N}^3$ is the flux network to which we apply the simplification algorithm. If we remove the species X first from V_3, we obtain a flux network representing the reaction network N''_X, and from that we get a flux network representing N''_{XYZ} by eliminating Y and Z (in any order). This flux network has only one flux vector which*

is $1^2 2^2 3^2 = (2(123))^{\mathbb{N}^3}$. *If we remove Y first we obtain a flux network representing* N''_Y, *and from that a flux network representing* N''_{YXZ} *by removing X and Z. The latter flux network is the singleton with the flux vector* 123.

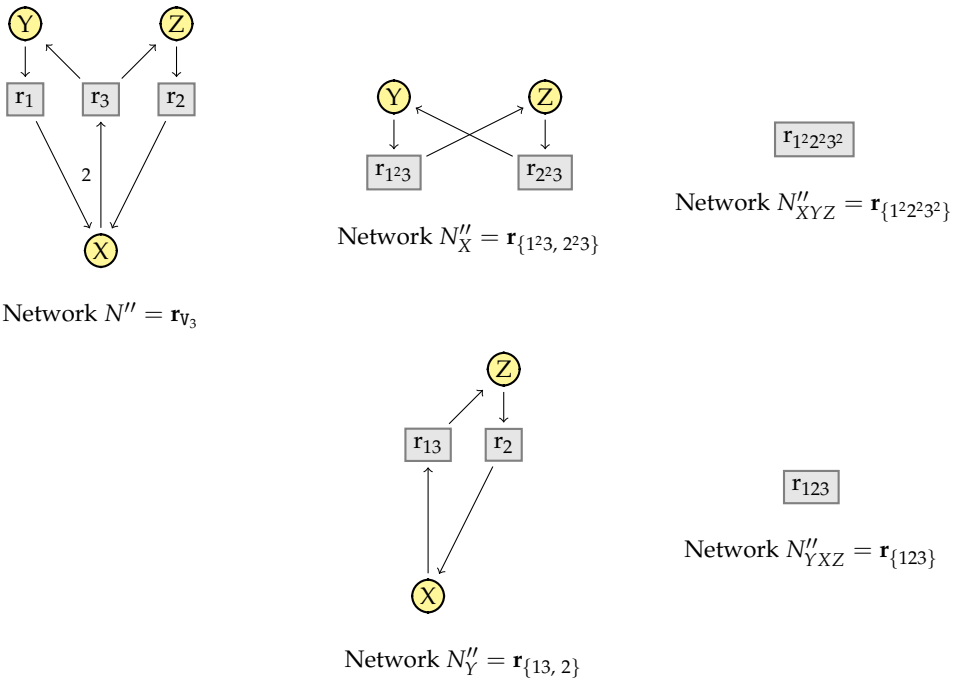


Figure 10. Elimination of intermediate species from flux networks in different orders is not confluent without factorization.

What is needed is a rule for the factorization of scalar multiples $(kv)^{\mathbb{N}^n}$ of vector v. This is done by the rule (F-FACT) in Figure 9, which, in the previous example, allows to simplify N''_{XYZ} into N''_{YXZ}.

We first note a consequence of the Folklore Theorem 1 and the following Lemma that is equally well known.

Lemma 2. *Let* $v, v' \in \mathbb{N}^n$ *be two elementary modes of the same matrix S. If* $supp(v) = supp(v')$ *then* $v = v'$.

Proof. Suppose that $supp(v) = supp(v')$. Write $v = (n_1, \ldots, n_n)$ and $v' = (n'_1, \ldots, n'_n)$. Let $i \in supp(v)$ be such that n_i/n'_i is maximal. Without loss of generality we can assume that $n_i/n'_i \geq 1$ since otherwise, we can exchange v and v'. Consider the vector of integers $w = n_i v' - n'_i v$. For any $j \in supp(v)$ we have:

$$n_i n'_j - n'_i n_j = n'_i n'_j (n_i/n'_i - n_j/n'_j) \geq 0.$$

Therefore, $w \in \mathbb{N}^n$, and thus $w \in ker_+(S)$. Furthermore, $i \notin supp(w)$ so that $supp(w) \subsetneq supp(v)$. Since v is an elementary mode of S this implies that $w = 0$, and thus $n_i v' = n'_i v$. Without loss of generality we can assume that n_i and n'_i have no common prime factors. If $n_i = n'_i = 1$ we are done. Otherwise $n_i \geq 2$ since $n_i/n'_i \geq 1$. Thus v can be factorized by n_i, contradiction. □

Corollary 1. *Let S be an* $m \times n$ *matrix of integers and* $E \subseteq \mathbb{N}^n$ *the set of all elementary modes of S. For any set* $E' \subseteq \mathbb{N}^n$ *such that* $cone(E') = ker_+(S)$, *if* E' *is irreducible by* $\Rrightarrow_{\text{F-FACT}}$ *and* $\Rrightarrow_{\text{F-DEP}}$ *then* $E' = E$.

Proof. By Theorem 1, we have $cone(E) = ker_+(S) = cone(E')$.

We first show that $E \subseteq E'$. Let $v \in E$. Since $v \neq 0^{\mathbb{N}^n}$ and $E \subseteq cone(E')$, v is of the form $v = n^1 v^1 + \ldots + n^k v^k$ for some $k \geq 1$, factorized $v^i \in E' \setminus \{0^{\mathbb{N}^n}\}$ and $n^i \in \mathbb{N} \setminus \{0\}$. Since all n^i

and v^i are positive, it follows that $supp(v^i) \subseteq supp(v)$ for all $1 \leq i \leq k$. Consider $i = 1$. Since v is an elementary mode and $v^1 \in ker_+(S)$, this implies that $supp(v^1) = supp(v)$. Since v^1 is factorized, and a member of $ker_+(S)$ with minimal support, it is also an elementary mode of S. Lemma 2 thus implies that $v^1 = v$, and so $v \in E'$. (It also follows that $k = n^1 = 1$).

We next show that $E' \subseteq E$. Let $v \in E'$. Since $E' \subseteq cone(E') = cone(E)$, vector v has the form $v = n^1v^1 + \ldots + n^kv^k$ for some $v^i \in E$. Since $E \subseteq E'$ and E' is closed by rule (F-DEP) it follows that $k = 1$. Hence, $v = n^1v^1$. Since E' is closed by rule (F-FACT) it follows that $n^1 = 1$. Hence $v = v^1 \in E$. □

5.4. Proving Confluence via Elementary Modes

Given a tuple of initial reactions $\mathbf{r}$ of size n and a set of intermediates $\mathcal{I} \subseteq Spec$ as parameters, we obtain a simplification relation on flux networks:

$$\Rrightarrow_{\text{F}} =_{\text{df}} (\Rrightarrow_{\text{F-INTER}} \cup \Rrightarrow_{\text{F-DEP}} \cup \Rrightarrow_{\text{F-FACT}}).$$

We now show that this relation is confluent for all possible choices of the parameters. The proof is by reduction to the Corollary 1 of the folklore Theorem 1 on elementary modes. We start with an fundamental property of the diamond operator $\diamond_X$, that we formulate in a sufficiently general manner so that is can be reused later on.

Lemma 3 (Diamond). *Let $(\mathcal{G}, +^{\mathcal{G}}, 0^{\mathcal{G}}, .^{\mathcal{G}}, 1^{\mathcal{G}})$ be a commutative semi-ring and $h : \mathbb{N}^n \to \mathcal{G}$ a semi-group homomorphism with respect to addition. Given a tuple $(v_1, \ldots, v_k)$ of vectors in $\mathbb{N}^n$, a tuple $(g_1, \ldots, g_k)$ of elements of $\mathcal{G}$, and a species $X \in Spec$, we define:*

$$\begin{array}{rclcrcl} P & = & \{p \in \{1...k\} \mid \boldsymbol{r}_{v_p}(X) > 0\}, & & prod & = & (\sum_{p\in P} \boldsymbol{r}_{v_p}(X) g_p)^{\mathcal{G}}, \\ C & = & \{c \in \{1...k\} \mid \boldsymbol{r}_{v_c}(X) < 0\}, & & cons & = & (\sum_{c\in C} -\boldsymbol{r}_{v_c}(X) g_c)^{\mathcal{G}} \end{array}$$

It then holds that:

$$\textstyle\sum^{\mathcal{G}}_{p\in P} \sum^{\mathcal{G}}_{c\in C} g_p .^{\mathcal{G}} g_c .^{\mathcal{G}} h(v_p \diamond_X v_c) \;=\; \sum^{\mathcal{G}}_{p\in P} g_p .^{\mathcal{G}} cons .^{\mathcal{G}} h(v_p) +^{\mathcal{G}} \sum^{\mathcal{G}}_{c\in C} g_c .^{\mathcal{G}} prod .^{\mathcal{G}} h(v_c)$$

Proof. We use some elementary rules of commutative semi-rings to distribute and factorize the sums contained in the definition of the diamond:

$$\begin{array}{rcl} & & \sum^{\mathcal{G}}_{p\in P} \sum_{Gc\in C} g_p .^{\mathcal{G}} g_c .^{\mathcal{G}} h(v_p \diamond_X v_c) \\ & = & \sum^{\mathcal{G}}_{p\in P} \sum^{\mathcal{G}}_{c\in C} g_p .^{\mathcal{G}} g_c .^{\mathcal{G}} ((-\mathbf{r}_{v_c}(X) h(v_p))^{\mathcal{G}} +^{\mathcal{G}} (\mathbf{r}_{v_p}(X) h(v_c))^{\mathcal{G}}) \\ & = & \sum^{\mathcal{G}}_{p\in P} \sum^{\mathcal{G}}_{c\in C} g_p .^{\mathcal{G}} g_c .^{\mathcal{G}} (-\mathbf{r}_{v_c}(X) h(v_p))^{\mathcal{G}} +^{\mathcal{G}} \sum^{\mathcal{G}}_{p\in P} \sum^{\mathcal{G}}_{c\in C} g_p .^{\mathcal{G}} g_c .^{\mathcal{G}} (\mathbf{r}_{v_p}(X) h(v_c))^{\mathcal{G}} \\ & = & \sum^{\mathcal{G}}_{p\in P} g_p .^{\mathcal{G}} (\sum_{c\in C} -\mathbf{r}_{v_c}(X) g_c)^{\mathcal{G}} .^{\mathcal{G}} h(v_p) +^{\mathcal{G}} \sum^{\mathcal{G}}_{c\in C} g_c .^{\mathcal{G}} (\sum_{p\in P} \mathbf{r}_{v_p}(X) g_p)^{\mathcal{G}} .^{\mathcal{G}} h(v_c) \\ & = & \sum^{\mathcal{G}}_{p\in P} g_p .^{\mathcal{G}} cons .^{\mathcal{G}} h(v_p) +^{\mathcal{G}} \sum^{\mathcal{G}}_{c\in C} g_c .^{\mathcal{G}} prod .^{\mathcal{G}} h(v_c) \end{array}$$

□

Our next objective is to show that the simplification preserves the invariants, when relativised to $\mathbf{r}$. For any flux network V, we therefore define the set of relatived invariants of V as follows:

$$inv_{\mathbf{r}}(V) = \{(n_1v_1 + \ldots + n_kv_k)^{\mathbb{N}^n}\mathbf{r} \mid n_1\mathbf{r}_{v_1} + \ldots + n_k\mathbf{r}_{v_k} \in inv(\mathbf{r}_V)\}.$$

For $\mathtt{V}_n = \{(1,0,\ldots,0), \ldots, (0,\ldots,0,1)\} \subseteq \mathbb{N}^n$ with the vectors ordered in the way they are enumerated, note that we have $\mathbf{r}_{\mathtt{V}_n} = \mathtt{N}_{\mathbf{r}}$ and $inv_{\mathbf{r}}(\mathtt{V}_n) = inv(\mathtt{N}_{\mathbf{r}})$. We next show that such relativised invariants are preserved by the simplification of flux networks.

Lemma 4. *If $V \Rrightarrow_F V'$ then $inv_{\boldsymbol{r}}(V) = inv_{\boldsymbol{r}}(V')$.*

Proof. We assume $V \Rrightarrow_{\text{F}} V'$ and first show the inclusion $inv_{\mathbf{r}}(V) \subseteq inv_{\mathbf{r}}(V')$.

Let $n_1 v_1 + \ldots + n_k v_k \in inv_{\mathbf{r}}(V)$. Then $n_1 \mathbf{r}_{v_1} + \ldots + n_k \mathbf{r}_{v_k} \in inv(\mathbf{r}_V)$. This means $(n_1 \mathbf{r}_{v_1} + \ldots + n_k \mathbf{r}_{v_k})^{\mathcal{R}} = 0^{\mathcal{R}}$. Since $\Rrightarrow_{\text{F}}$ is the union $\Rrightarrow_{\text{F-FACT}} \cup \Rrightarrow_{\text{F-DEP}} \cup \Rrightarrow_{\text{F-INTER}}$, three cases are to be considered.

Case $V \Rrightarrow_{\text{F-FACT}} V'$**.** Suppose that (F-FACT) replaces vector v_1 by vector v_1' so that $v_1 = k' v_1'$ for some $k' \neq 0$. Hence $n_1 k' \mathbf{r}_{v_1'} + n_2 \mathbf{r}_{v_2} + \ldots + n_k \mathbf{r}_{v_k} \in inv(\mathbf{r}_{V'})$. And thus, $(n_1 k' v_1' + n_2 v_2 + \ldots + n_k v_k)^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V')$, which is equivalent to $(n_1 v_1 + \ldots + n_k v_k)^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V')$ as required.

Case $V \Rrightarrow_{\text{F-DEP}} V'$**.** By rule (F-DEP) there exist $k \in \mathbb{N}$, $\mathbf{v} \in V^k$ and $v \in \mathbb{N}^k$ such that $V = V' \uplus \{\mathbf{v}_v\}$. If all v_i are distinct from $\mathbf{v}_v$ then trivially $n_1 \mathbf{r}_{v_1} + \ldots + n_k \mathbf{r}_{v_k} \in inv(\mathbf{r}_{V'})$. Otherwise, we can assume without loss of generality that $v_1 = \mathbf{v}_v$ with v and $\mathbf{v}$ as in rule (F-DEP). Suppose that these have the forms $v = (m_1, \ldots, m_l)$ and $\mathbf{v} = (w_1, \ldots, w_l)$. Since $\mathbf{r}_{\mathbf{v}_v} = (m_1 \mathbf{r}_{w_1} + \ldots + m_l \mathbf{r}_{w_l})^{\mathcal{R}}$, it follows that:

$$n_1 m_1 \mathbf{r}_{w_1} + \ldots + n_1 m_l \mathbf{r}_{w_l} + n_2 \mathbf{r}_{v_2} + \ldots + n_k \mathbf{r}_{v_k} \in inv(\mathbf{r}_{V'}).$$

This yields $(n_1 m_1 w_1 + \ldots + n_1 m_l w_l + n_2 v_2 + \ldots + n_k v_k)^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V')$. Since $v_1 = \mathbf{v}_v = (m_1 w_1 + \ldots + m_l w_l)^{\mathbb{N}^n}$ this is is equivalent to $(n_1 v_1 + \ldots + n_k v_k)^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V')$ as required.

Case $V \Rrightarrow_{\text{F-INTER}} V'$**.** Suppose that the intermediate species $X \in \mathcal{I}$ was eliminated thereby. Recall that $\sum_{i=1}^{k} n_i \mathbf{r}_{v_i} \in inv(\mathbf{r}_V)$. We can assume without loss of generality that $n_i \neq 0$ for all $1 \leq i \leq k$. Let P, C, *prod*, and *cons* be as introduced in the Diamond Lemma 3, where $\mathcal{G} = \mathbb{N}^n$, homomorphism h the identity on $\mathbb{N}^n$, and $g_i = n_i$ for all $1 \leq i \leq k$. The lemma then yields:

$$(\textstyle\sum_{p \in P} \sum_{c \in C} n_p n_c \ (v_p \diamond_X v_c))^{\mathbb{N}^n} \ = \ (\sum_{p \in P} n_p \ cons \ v_p + \sum_{c \in C} n_c \ prod \ v_c)^{\mathbb{N}^n} \ .$$

Since $(\sum_{i=1}^{k} n_i \mathbf{r}_{v_i})^{\mathcal{R}} = 0^{\mathcal{R}}$ it follows that $prod = cons$. Furthermore, $prod \neq 0$ since otherwise $P = C = \emptyset$ so that (F-INTER) could not be applied. Since $cons = prod$, this tuple is equal to $prod(\sum_{p \in P} n_p v_p + \sum_{c \in C} n_c v_c)^{\mathbb{N}^n}$. With $M = \{m \in \{1...k\} \mid \mathbf{r}_{v_m}(X) = 0\}$ we get:

$$(\textstyle\sum_{p \in P} \sum_{q \in C} n_p n_c (v_p \diamond_X v_c) + \sum_{m \in M} prod \ n_m v_m)^{\mathbb{N}^n} \mathbf{r} \ = \ (\sum_{i=1}^{k} prod \ n_i v_i)^{\mathbb{N}^n} \mathbf{r} \ .$$

This multiset is an invariant, since $(\sum_{i=1}^{k} n_i \mathbf{r}_{v_i})^{\mathcal{R}} = 0^{\mathcal{R}}$. It follows that:

$$(\textstyle\sum_{p \in P} \sum_{c \in C} n_p n_c (v_p \diamond_X v_c) + \sum_{m \in M} n_m \ prod \ v_m)^{\mathbb{N}^n} \mathbf{r} \in inv(\mathbf{r}_{V'}) \ .$$

This implies $(\sum_{i=1}^{k} prod \ n_i v_i)^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V')$. Since $prod \neq 0$ and since $inv_{\mathbf{r}}(V')$ is closed by factorization with nonzero factors, it follows that $\sum_{i=1}^{k} n_i \mathbf{r}_{v_i} \in inv_{\mathbf{r}}(V')$ as required.

The proof of the inverse inclusion $inv_{\mathbf{r}}(V) \supseteq inv_{\mathbf{r}}(V')$ differs in that the Diamond Lemma is not needed. Let $n_1 v_1' + \ldots + n_k v_k' \in inv_{\mathbf{r}}(V')$. Then $n_1 \mathbf{r}_{v_1'} + \ldots + n_k \mathbf{r}_{v_k'} \in inv(\mathbf{r}_{V'})$. This means $(n_1 \mathbf{r}_{v_1'} + \ldots + n_k \mathbf{r}_{v_k'})^{\mathcal{R}} = 0^{\mathcal{R}}$. We distinguish three cases depending on which rule was applied:

Case $V \Rrightarrow_{\text{F-FACT}} V'$**.** Suppose that (F-FACT) replaces vector v_1 by vector v_1' so that $v_1 = k' v_1'$ for some $k' \neq 0$. Since $(k' n_1 \mathbf{r}_{v_1'} + \ldots + k' n_k \mathbf{r}_{v_k'})^{\mathcal{R}} = 0^{\mathcal{R}}$ we have $n_1 \mathbf{r}_{v_1} + n_2 k' \mathbf{r}_{v_2'} + \ldots + n_k k' \mathbf{r}_{v_k'} \in inv(\mathbf{r}_V)$. And thus, $(n_1 v_1 + n_2 k' v_2' + \ldots + n_k k' v_k')^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V)$, which is equivalent to $(n_1 k' v_1' + \ldots + n_k k' v_k')^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V)$, and thus $(n_1 v_1' + \ldots + n_k v_k')^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V)$ as required.

Case $V \Rrightarrow_{\text{F-DEP}} V'$**.** By rule (F-DEP) there exist $k \in \mathbb{N}$, $\mathbf{v} \in V^k$ and $v \in \mathbb{N}^k$ such that $V = V' \uplus \{\mathbf{v}_v\}$. If all v_i' are distinct from $\mathbf{v}_v$ then trivially $n_1 \mathbf{r}_{v_1'} + \ldots + n_k \mathbf{r}_{v_k'} \in inv(\mathbf{r}_V)$. Otherwise, we can assume without loss of generality that $v_1' = \mathbf{v}_v$ with v and $\mathbf{v}$ as in the rule. Suppose that these have the forms $v = (m_1, \ldots, m_l)$ and $\mathbf{v} = (w_1, \ldots, w_l)$. Since $\mathbf{r}_{\mathbf{v}_v} = (m_1 \mathbf{r}_{w_1} + \ldots + m_l \mathbf{r}_{w_l})^{\mathcal{R}}$, it follows that:

$$n_1 m_1 \mathbf{r}_{w_1} + \ldots + n_1 m_l \mathbf{r}_{w_l} + n_2 \mathbf{r}_{v_2'} + \ldots + n_k \mathbf{r}_{v_k'} \in inv(\mathbf{r}_V).$$

This yields $(n_1 m_1 w_1 + \ldots + n_1 m_l w_l + n_2 v_2' + \ldots + n_k v_k')^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V)$. Since $v_1' = \mathbf{v}_v = (m_1 w_1 + \ldots + m_l w_l)^{\mathbb{N}^n}$ this is is equivalent to $(n_1 v_1' + \ldots + n_k v_k')^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V)$ as required.

Case $V \Rrightarrow_{\text{F-INTER}} V'$**.** Suppose that the intermediate species $X \in \mathcal{I}$ was eliminated thereby. We recall that $\sum_{i=1}^{k} n_i \mathbf{r}_{v_i'} \in inv(\mathbf{r}_{V'})$. Without loss of generality, we can assume that all elements of V' occur exactly once in this sum. Let $V = \{v_1, \ldots, v_l\}$, $P = \{p \mid \mathbf{r}_{v_p}(X) > 0\}$, $C = \{c \mid \mathbf{r}_{v_c}(X) < 0\}$, and $M = \{m \mid \mathbf{r}_{v_m}(X) = 0\}$. If $v_i' = v_p \diamond_X v_c$ for $p \in P$ and $c \in C$, we note $o_{pc} = n_i$. Otherwise, if $v_i' = v_m$ with $m \in M$, we note $o_m = n_i$. By the rule (F-INTER) we have:

$$\begin{aligned}(\textstyle\sum_{i=1}^{k} n_i v_i')^{\mathbb{N}^n} &= (\textstyle\sum_{p\in P}\sum_{c\in C} o_{pc} v_p \diamond_X v_c + \sum_{m\in M} o_m v_m)^{\mathbb{N}^n} \\ &= \left(\textstyle\sum_{p\in P}(\sum_{c\in C} o_{pc}\mathbf{r}_{v_c}(X))v_P + \sum_{c\in C}(\sum_{p\in P} o_{pc}\mathbf{r}_{v_p}(X))v_C + \sum_{m\in M} o_m v_m\right)^{\mathbb{N}^n}\end{aligned}$$

Hence $(\sum_{i=1}^{k} n_i v_i')^{\mathbb{N}^n} \mathbf{r} \in inv_{\mathbf{r}}(V)$.

□

We start the reminder of the proof with the case where all species are intermediates so that $\mathcal{I} = Spec$.

Lemma 5. *If $\mathcal{I} = Spec$ and V is irreducible by $\Rrightarrow_{\text{F-INTER}}$ then $\{\boldsymbol{v}_v \boldsymbol{r} \mid \boldsymbol{v} \in V^k,\ v \in \mathbb{N}^k,\ k \in \mathbb{N}\} = inv_{\boldsymbol{r}}(V)$.*

Proof. Given that V is irreducible by $\Rrightarrow_{\text{F-INTER}}$, all intermediates species must be eliminated in all reactions of $\mathbf{r}_V$. Since $\mathcal{I} = Spec$ this implies that all species are eliminated in all reactions of $\mathbf{r}_V$, so for all $v \in \mathbf{r}_V$ it follows that $\mathbf{r}_v = 0^{\mathcal{R}}$. Thus for any $v \in V$ and $\mathbf{v} \in V^k$ we have $(\mathbf{v}_v)^{\mathcal{R}} = 0^{\mathcal{R}}$, so that $\mathbf{v}_v \in inv_{\mathbf{r}}(V)$. Hence $\{\mathbf{v}_v \mathbf{r} \mid \mathbf{v} \in V^k,\ v \in \mathbb{N}^k,\ k \in \mathbb{N}\} \subseteq inv_{\mathbf{r}}(V)$. The inverse inclusion holds trivially. □

Proposition 1. *Let $\mathcal{I} = Spec$ and $\mathtt{V}_n \Rrightarrow_{\text{F}}^{*} V$ such that V irreducible for $\Rrightarrow_{\text{F-INTER}}$. Then $cone(V) = ker_+(S)$, where S is the stoichiometry matrix of $\mathtt{N}_{\mathbf{r}}$.*

Proof. By Lemmas 4 and 5 we have: $\{\mathbf{v}_v \mathbf{r} \mid \mathbf{v} \in V^k,\ v \in \mathbb{N}^k,\ k \in \mathbb{N}\} = inv_{\mathbf{r}}(V) = inv_{\mathbf{r}}(\mathtt{V}_n) = inv(\mathtt{N}_{\mathbf{r}})$. This yields $\{\mathbf{v}_v \mid \mathbf{v} \in V^k,\ v \in \mathbb{N}^k,\ k \in \mathbb{N}\} = ker_+(S)$, i.e., $cone(V) = ker_+(S)$. □

Theorem 2. *Consider the simplification relation for flux networks $\Rrightarrow_{\text{F}}$ that is parametrised by $\mathcal{I} = Spec$ and a tuple of initial reactions $\mathbf{r}$. If $\mathtt{V}_n \Rrightarrow_{\text{F}}^{*} V$ for some flux network V that is irreducible for $\Rrightarrow_{\text{F}}$, then $V = E$, where E is the set of elementary modes of the stoichiometry matrix of $\mathtt{N}_{\mathbf{r}}$.*

Proof. From Proposition 1 it follows that $cone(V) = ker_+(S)$ where S is the stoichiometry matrix of $\mathtt{N}_{\mathbf{r}}$. Furthermore, V is irreducible with respect to $\Rrightarrow_{\text{F-FACT}} \cup \Rrightarrow_{\text{F-DEP}}$, so that Corollary 1 implies $V = E$. □

Corollary 2. *The simplification relation $\Rrightarrow_{\text{F}}$ restricted to flux networks in the set $\{V \mid \mathtt{V}_n \Rrightarrow_{\text{F}}^{*} V\}$ is confluent.*

Proof. We notice that $\Rrightarrow_{\text{F}}$ is terminating, since (F-INTER) reduces the number of intermediate species $X \in \mathcal{I}$ for which there exists a vector v such that $\mathbf{r}_v(X) \neq 0$, (F-DEP) reduces the number of vectors in the set, and (F-FACT) reduces the norm of one of the vectors.

We first consider the case $Spec = \mathcal{I}$. Let V be such that $\mathtt{V}_n \Rrightarrow_{\text{F}}^{*} V$, where $\Rrightarrow_{\text{F}}$ is parametrised by $\mathcal{I}$ and a tuple $\mathbf{r}$ of initial reactions. Suppose that $V \Rrightarrow_{\text{F}}^{*} V_1$ and $V \Rrightarrow_{\text{F}} V_2$. Since $\Rrightarrow_{\text{F}}$ is terminating there exist V_1' and V_2' that are irreducible with $\Rrightarrow_{\text{F}}$ such that $V_1 \Rrightarrow_{\text{F}}^{*} V_1'$ and $V_1 \Rrightarrow_{\text{F}}^{*} V_2'$. Theorem 2 proves that $V_1' = E = V_2'$, where E is the set of elementary modes of the stoichiometry matrix of $\mathbf{r}_{\mathtt{V}_n} = \mathtt{N}_{\mathbf{r}}$.

We next reduce the general case where $Spec \subseteq \mathcal{I}$ to the case $Spec = \mathcal{I}$. We define $\mathbf{r}_{|\mathcal{I}}$ by restricting all reactions in the tuple $\mathbf{r}$ to $\mathcal{I}$, i.e., if $\mathbf{r} = (r_1, \ldots, r_n)$ then $\mathbf{r}_{|\mathcal{I}} = (r_{1|\mathcal{I}}, \ldots, r_{n|\mathcal{I}})$. We then observe that the relation $\Rrightarrow_{\text{F}}$ with respect to $\mathbf{r}$ coincides with the relation $\Rrightarrow_{\text{F}}$ with respect to $\mathbf{r}_{|\mathcal{I}}$. Hence the confluence result from the case $\mathcal{I} = Spec$ can be applied. □

As shown by Theorem 2, the exhaustive simplification of flux networks V with $\Rrightarrow_{\text{F}}$ can be used to compute the set of elementary modes of the stoichiometry matrix of the reaction network $\mathbf{r}_V$.

Interestingly, this algorithm is essentially the same as the double description method, as recalled for instance in [21]. The correspondence comes from the fact that any reaction network can be identified with its stoichiometry matrix, so that the algorithm can be formulated either for the one or the other representation. Still there is a minor difference between this algorithm and the one in [21]. The algorithm presented here is slightly more flexible, in that the rule (F-DEP) can be applied at any stage of the simplification while in the double description method as described in [21], the rule (F-DEP) is applied at the same time as the rule (F-INTER). However, as we have shown with the confluence Theorem 2, this additional freedom in the application order of the rules does not affect the final result.

6. Reaction Networks with Deterministic Semantics

We now consider reactions with kinetic expressions, and recall some basic definitions. We first define expressions and networks with kinetics. Then we recall how to associate a system of equations to a reaction network. Finally we use this system of equations to define the deterministic semantics of reaction networks.

6.1. Kinetic Expressions

We now define a class of kinetic expressions. Their syntax is the same as that of arithmetic expressions, by their semantics is by interpretation as functions of type $\mathbb{R}_+ \to \mathbb{R}$,

Let *Param* be a set of parameters of type $\mathbb{R}_+$. As set of variables of type $\mathbb{R}_+ \to \mathbb{R}_+$, we will use the set *Spec*. A variable $A \in Spec$ is intended to represent the temporal evolution of the concentration of A over time.

We define the set of *expressions Expr* by the terms with the abstract syntax in Figure 11. Expressions describe functions of type $\mathbb{R}_+$ to $\mathbb{R}$. They are built from species A of type $\mathbb{R}_+$ to $\mathbb{R}_+$, and constant functions defined by parameters $k \in \mathbb{R}_+$, constants $c \in \mathbb{R}$, and expressions $e(0)$, standing for the value of e at time 0. Beside of these, expressions can be constructed by addition, subtraction, multiplication, and division. For convenience, we will use parenthesis (e) whenever the priority of the operators might not be clear. For any species e we denote by $pSpec(e)$ the subset of species that occur properly in e, that is outside of a sub-expression $e(0)$. So for instance $pSpec(B = A(0)) = \{B\}$.

Syntax

$$e \in Expr ::= A \mid k \mid c \mid e(0) \mid e + e' \mid ee' \mid 1/e \mid -e$$

Shortcuts

$$e/e' \quad =_{\text{df}} \quad e(1/e'), \qquad e - e' \quad =_{\text{df}} \quad e + (-e'), \qquad e^n \quad =_{\text{df}} \quad \underbrace{e \ldots e}_{n \text{ times}}.$$

Semantics

$$[\![k]\!]_\alpha = \left\{ \begin{array}{ccc} \mathbb{R}_+ & \to & \mathbb{R}_+ \\ t & \mapsto & \beta(k) \end{array} \right. \qquad [\![c]\!]_\alpha = \left\{ \begin{array}{ccc} \mathbb{R}_+ & \to & \mathbb{R}_+ \\ t & \mapsto & c \end{array} \right.$$

$$[\![A]\!]_\alpha = \alpha(x) \qquad [\![-e]\!]_\alpha = -^{\mathbb{R}}[\![e]\!]_\alpha$$

$$[\![e(0)]\!]_\alpha = \left\{ \begin{array}{ccc} \mathbb{R}_+ & \to & \mathbb{R}_+ \\ t & \mapsto & [\![e]\!]_\alpha(0) \end{array} \right. \qquad [\![e + e']\!]_\alpha = [\![e]\!]_\alpha +^{\mathbb{R}} [\![e']\!]_\alpha$$

$$[\![1/e]\!]_\alpha = \left\{ \begin{array}{ll} 1/^{\mathbb{R}}[\![e]\!]_\alpha & \text{if } \forall t \in \mathbb{R}_+.\ [\![e]\!]_\alpha(t) \neq^{\mathbb{R}} 0 \\ \bot & \text{otherwise} \end{array} \right. \qquad [\![ee']\!]_\alpha = [\![e]\!]_\alpha *^{\mathbb{R}} [\![e']\!]_\alpha$$

Figure 11. Expressions where $A \in Spec$, $k \in Param$, $c \in \mathbb{R}$, and $n \in \mathbb{N}$.

The semantics of expressions is parametrised by a function $\beta : Param \to \mathbb{R}_+$ that interprets all parameters as positive real numbers. In order to simplify the notation, we assume that β is fixed, but notice that our simplification algorithms will be correct for any interpretation β.

The value of an expression $[\![e]\!]_\alpha \in (\mathbb{R}_+ \to \mathbb{R}_+) \cup \{\bot\}$ is specified in Figure 11 for any variable assignment $\alpha : Vars \to (\mathbb{R}_+ \to \mathbb{R}_+)$. It may either be a function of type $\mathbb{R}_+ \to \mathbb{R}$ or undefined $\bot$. The latter is necessary for the interpretation of $[\![1/e]\!]_\alpha$, which is defined only if $[\![e]\!]_\alpha(t) \neq 0$ for any time point t. We call an expression e nonnegative if $e \geq 0$ is valid, i.e., if for all non-negative assignment α and all time points $t \in \mathbb{R}_+$, we have $[\![e]\!]_\alpha(t) \geq 0$.

Definition 8. *A kinetic expression is a nonnegative expression $e \in Expr$.*

6.2. Constrained Flux Networks

The next objective is to add kinetic expressions to reactions and flux networks. Furthermore, we need to be able to express constraints about these kinetic expressions in order to express partial steady state hypothesis and conservation laws. This will lead us to the notion of constrained flux networks. A reaction with kinetics expressions is a pair $r; e$ where r is a reaction without kinetics and e is a kinetic expression. As before we now use flux reaction to represent a reaction but now with a kinetic expression.

Definition 9. *An n-ary flux reaction with kinetic expression is a pair $v; e$ composed of a vector $v \in \mathbb{N}^n$ and a kinetic expression $e \in Expr$. Given a tuple of reactions $\boldsymbol{r} = (r_1, \ldots, r_n)$, the flux reaction $v; e$ represents the reaction $\boldsymbol{r}_v; e$.*

The set $\mathcal{C}$ of *constraints on kinetic functions* is defined in Figure 12. A constraint $C \in \mathcal{C}$ is a conjunction of atomic constraints. The first kind is an equation $e = e'$ stating that the expressions e and e' must have the same value but different from $\bot$. The atomic constraint $cst(e)$ requires that e is a constant function, $e \neq 0$ that e may never becomes equal to zero, and $e \geq 0$ that e is always non-negative. More formally, we define in Figure 12 the interpretation $[\![C]\!]_\alpha \in \mathbb{B} \cup \{\bot\}$ of a constraint C for a given variable assignment α, where $\mathbb{B} = \{true, false\}$ is the set of boolean values.

Syntax

$$C \in \mathcal{C} ::= e = e' \mid e \neq 0 \mid e \geq 0 \mid cst(e) \mid C \wedge C' \mid true$$

Shortcuts

$$e > 0 =_{\text{df}} e \geq 0 \wedge e \neq 0.$$

Semantics

$$[\![e = e']\!]_\alpha = \begin{cases} [\![e]\!]_\alpha = [\![e']\!]_\alpha & \text{if } [\![e]\!]_\alpha, [\![e']\!]_\alpha \in \mathbb{R}_+ \to \mathbb{R} \\ \bot & \text{otherwise} \end{cases} \qquad [\![cst(e)]\!]_\alpha = \exists c. \forall t.\ [\![e]\!]_\alpha(t) =^{\mathbb{R}} c$$

$$[\![e \neq 0]\!]_\alpha = \forall t \in \mathbb{R}_+.\ [\![e]\!]_\alpha(t) \neq^{\mathbb{R}} 0 \qquad [\![C \wedge C']\!]_\alpha = [\![C]\!]_\alpha \wedge^{\mathbb{B}} [\![C']\!]_\alpha$$

$$[\![e \geq 0]\!]_\alpha = \forall t \in \mathbb{R}_+.\ [\![e]\!]_\alpha(t) \geq^{\mathbb{R}} 0 \qquad [\![true]\!]_\alpha = true$$

Figure 12. Constraints on kinetic functions.

Definition 10. *An n-ary constrained flux network is a pair $W = V\&C$ where V is a set of n-ary flux reactions with kinetic expressions and C a constraint.*

Let $W = V\&C$ be a constrained flux network. We denote by $Expr(W)$ the set of kinetic expressions e such that $v; e \in V$ or such that e occurs in the constraint C. We set:

$$pSpec(W) = \{A \mid v; e \in V,\ \mathbf{r}_v(A) \neq 0 \text{ or } A \in pSpec(e)\} \cup pSpec(C).$$

6.3. Systems of Constrained Equations with ODEs

We now recall how to assign systems of equations to constrained flux networks. Note that systems constrained equations may contain both constraints and *ordinary differential equations* (ODEs) in particular.

The set of systems of constrained equations is defined in Figure 13. They are conjunctions of constraints C and ODEs $\dot{A} = e$ where $A \in Spec$ and $e \in Expr$. Note that the constraints may subsume the non-differential arithmetic equations $e = e'$. We denote by $Spec(E)$ the set of (free) variables occurring in E, and by $Expr(E)$ the set of expressions contained in E.

Syntax

$$E ::= \dot{A} = e \mid C \mid E \wedge E'$$

Semantics

$$[\![\dot{A} = e]\!]_\alpha = \begin{cases} \frac{d[\![A]\!]_\alpha(t)}{dt} =^{\mathbb{R}_+ \to \mathbb{R}} [\![e]\!]_\alpha & \text{if for all } t \in \mathbb{R}_+ \colon [\![e]\!]_\alpha(t) \neq \bot \\ \bot & \text{otherwise} \end{cases}$$

$[\![C]\!]_\alpha = \ldots$ see Figure 12

$[\![E \wedge E']\!]_\alpha = [\![E]\!]_\alpha \wedge^{\mathbb{B}} [\![E']\!]_\alpha$.

Figure 13. Systems of constrained equations with ODEs.

The denotation of a system of constrained equations E is a value in $[\![E]\!]_\alpha \in \mathbb{B} \cup \{\bot\}$ as defined in Figure 13. The set of solutions of $sol(E)$ is the set of assignments $\alpha : Spec \to \mathbb{R}_+$ that make E true, i.e.,

$$sol(E) = \{\alpha \mid [\![E]\!]_\alpha = true\}.$$

We say that a constrained equation E logically implies another E' and write $E \models E'$ if $sol(E) \subseteq sol(E')$. For instance, $true \models kA + kB = k(A + B)$, $e \neq 0 \models e/e = 1$, $cst(A) \models \dot{A} = 0$.

Definition 11. *Two constrained equation systems E and E' are called* logically equivalent, *denoted $E \models\!\!\models E'$, if they have the same solutions, i.e.,*

$$E \models\!\!\models E' \text{ iff } sol(E) = sol(E').$$

Clearly, $E \models\!\!\models E'$ if and only if E and E' logically imply each other, i.e., $E \models E'$ and $E' \models E$.

6.4. Deterministic Semantics

We assign to any constrained flux network $W = V\&C$ a system of constrained equations $E(W)$ in Figure 14. Note that $E(W)$ does depend on the tuple of initial reactions $\mathbf{r}$ and the set of intermediates $\mathcal{I}$. The system contains an ODE for any species A stating that the change of the concentration of A is equal to the sum of the rates $\mathbf{r}_v(A)e$ of the flux reactions $v; e \in V$. The factor $\mathbf{r}_v(A)$ makes the rate negative if A is consumed and positive if A is produced. It also takes care of the multiplicities of consumption and production. Finally, the constraint C of the constrained flux network is added to the system of constrained equations.

$$E(V\&C) = C \wedge \bigwedge_{A \in Spec} \dot{A} = \sum_{v;e \in V} \mathbf{r}_v(A)e$$

Figure 14. System of constrained equations of a constrained flux network $V\&C$.

Example 5. *We consider the flux network for the classical Michaelis-Menten example [10]. Its system of constrained equations is then represented in Figure 15. It contains four* ODEs *and two constant constraints.*

$$\begin{array}{lll} & \dot{S} = -k_1SE + k_2C & \wedge \quad \dot{E} = -k_1SE + (k_2 + k_3)C \\ \wedge & \dot{C} = k_1SE - (k_2 + k_3)C & \wedge \quad \dot{P} = k_3C \\ \wedge & cst(E) & \wedge \quad cst(C) \end{array}$$

Figure 15. System of constrained equations for Michaelis-Menten.

6.5. Contextual Equivalence

Two constrained flux networks W and W' are *non-contextually equivalent*, denoted $W \simeq W'$, if their systems of constrained equations are logically equivalent:

$$W \simeq W' \qquad \text{iff} \qquad E(W) \models\!\models E(W)'.$$

We now extend the definition to a contextual equivalence. The idea is that networks can be exchanged with equivalent networks in any context, without affecting the semantics. As contexts, we use flux networks themselves $W' = V'\&C'$. We define the combination of a network $W = V\&C$ and the context W' as follows:

$$W \mid W' =_{\text{df}} V \cup V'\&C \wedge C'.$$

We now assume a set of intermediate species $\mathcal{I} \subseteq Spec$ and call a context W' compatible if $pSpecs(W') \cap \mathcal{I} = \emptyset$.

Definition 12. *Two constrained flux networks W and W' are* (contextually) equivalent *if they have the same solutions in any compatible context, that is:*

$$W \sim W' \qquad \textit{iff} \qquad \forall W'' \textit{compatible}.\ W \mid W'' \simeq W' \mid W''.$$

We note that the definition of equivalence of constrained flux networks has two parameters: $\mathbf{r}$ and $\mathcal{I}$. The equivalence $\sim$ depends on the tuple of initial reactions $\mathbf{r}$, since the non-contextual equivalence relation $\simeq$ relies on the deterministic semantics of constrained flux networks, which in turn depends on $\mathbf{r}$. The equivalence relation $\sim$ also depends on the set of intermediates $\mathcal{I}$ since the notion of compatibility depends on it.

Our simplification algorithm will rewrite constrained flux networks up to logical equivalence of constraints. Therefore, we can hope for confluence only up to logical equivalence. More formally, we defined the similarity relation $\cong$ as the least equivalence relation on constrained flux networks that satisfies the following two inference rules for all C, C', V, e, e':

$$\frac{C \models e = e'}{\{v; e\} \cup V\&C \cong \{v; e'\} \cup V\&C}, \qquad \frac{C \models\!\models C'}{V\&C \cong V\&C'}.$$

The first rule states that expressions that are logically equivalent under the constraints of the constrained flux network can be replaced by each other. The second rule allows to exchange logically equivalent constraints by each other. Similar networks are trivially equivalent:

Lemma 6. *Similarity $W \cong W'$ implies contextual equivalence $W \sim W'$.*

Proof. Straightforward from the definitions. □

7. Simplification of Constrained Flux Networks

Our next objective is to simplify constrained flux networks by lifting the confluent simplification algorithm for flux networks to the case with kinetic expressions. This will require to impose partial steady state and linearity restrictions on the constrained flux networks, since otherwise, we would

not know how to remove intermediates from the constrained equations assigned to the constrained flux network.

7.1. Linear Steadiness of Intermediate Species

The following restriction will allow us to eliminate an intermediate species from the constraint equations of a constrained flux network.

Definition 13. *We say that a species $X \in \mathcal{I}$ is* linearly steady *in a constrained flux network $V\&C$ if it satisfies the following four conditions:*

Partial steady state: *the concentration of X is steady, i.e., $C \models cst(X)$.*
Linear consumption: *if a reaction in V consumes X then its kinetic expression is linear in X, that is: if $v; e \in V$ such that $\mathbf{r}_v(X) < 0$ then $C \models e = Xe'$ for some expression e' such that $X \notin pSpecs(e')$.*
Independent production: *if a reaction in V produces X then its kinetic expression does not contain X except for subexpressions $X(0)$: for any $v; e \in V$, if $\mathbf{r}_v(X) > 0$ then $X \notin pSpecs(e)$.*
Nonzero consumption: *the consumption of X is nonzero: $C \models \sum\{e \mid v; e \in V,\ \mathbf{r}_v(X) < 0\} \neq 0$.*

Suppose that X is linearly steady in $W = V\&C$. Since X is in *partial steady state*, we have $C \models cst(X)$ and hence $C \models \dot{X} = 0$. The constrained equations of W thus imply that the production and consumption of X are equal:

$$E(W) \models prod = cons \qquad \text{where} \qquad \begin{cases} prod = \sum\{\, \mathbf{r}_v(X)e \mid v; e \in V,\ \mathbf{r}_v(X) > 0\} \\ cons = \sum\{-\mathbf{r}_v(X)e \mid v; e \in V,\ \mathbf{r}_v(X) < 0\} \end{cases}$$

The *linear consumption* of X imposes that $C \models cons = Xe$ for some expression e such that $X \notin pSpecs(e)$. The *independent production* of X imposes that $X \notin pSpecs(prod)$. Because of *nonzero consumption*, we have:

$$E(W) \models X = X\frac{prod}{cons} = \frac{prod}{e}.$$

where the expression $\frac{prod}{e}$ does not contain the species X properly. Therefore, we can eliminate the variable X from the constrained equation $E(W)$ by substituting X by $\frac{prod}{e}$. This give us hope that we can also eliminate linearly steady intermediate species from the constrained flux networks too by adapting the rule (F-INTER) to kinetic expressions.

7.2. Simplification

We now lift the simplification rules for flux networks to constrained flux networks. The lifted rules are presented in Figure 16. They define the simplification relation for constrained flux networks:

$$\Rrightarrow_{\mathrm{C}} \quad =_{\mathrm{df}} \quad \Rrightarrow_{\text{C-INTER}} \cup \Rrightarrow_{\text{C-MOD}} \cup \Rrightarrow_{\text{C-DEP}} .$$

The first rule (C-INTER) eliminates a linearly steady intermediate species X, by merging any pair of reactions, so that the one produces and another consumes X. The rule can be applied only under the hypothesis that the constraints of the network imply that X is linearly steady, and so that X is in partial steady state in particular. It should also be noticed that the conditions on the initial value of X are preserved by the constraint $X(0) = X\frac{prod}{cons}$. As argued above, the linear steadiness of X implies that the latter is equivalent to some other expression that does not contain species X properly. So except for constraints on the initial value $X(0)$, the species X got removed from the constrained flux network. The rule also replaces X by $X(0)$ in all the kinetic expressions of reactions, in which X is used as a modifier, i.e., $v; e \in V$ such that $\mathbf{r}_v = 0$ and $X \in pSpecs(e)$. Furthermore, the same substitution is applied to the constraints of the flux network.

(C-INTER)

$$\frac{X \in \mathcal{I} \qquad X \text{ is linearly steady in } V\&C \qquad prod = \sum_{\{v;e \in V \mid \mathbf{r}_v(X) > 0\}} \mathbf{r}_v(X)e \qquad cons = \sum_{\{v;e \in V \mid \mathbf{r}_v(X) < 0\}} -\mathbf{r}_v(X)e}{V\&C \Rrightarrow_{\text{C-INTER}} \begin{cases} \{v \diamond_X v'; \frac{ee'}{cons} \mid v;e \in V,\ v';e' \in V,\ \mathbf{r}_v(X) > 0,\ \mathbf{r}_{v'}(X) < 0\} \\ \cup \{v;e[X := X(0)] \mid v;e \in V,\ \mathbf{r}_v(X) = 0\} \\ \&C[X := X(0)] \wedge X(0) = X\frac{prod}{cons} \end{cases}}$$

(C-MOD)

$$\frac{X \in \mathcal{I} \qquad \forall v;e \in V.\ \mathbf{r}_v(X) = 0}{V\&C \Rrightarrow_{\text{C-MOD}} \{v;e[X := X(0)] \mid v;e \in V\}\&C[X := X(0)]}$$

(C-DEP)

$$\frac{n_1, \ldots, n_k \in \mathbb{N} \qquad v = (\sum_{1 \leq i \leq k} n_i v_i)^{\mathbb{N}^n}}{\{v_i;e_i \mid 1 \leq i \leq k\} \cup \{v;e\} \cup V\&C \quad \Rrightarrow_{\text{C-DEP}} \quad \{v_i;e_i + n_i e \mid 1 \leq i \leq k\} \cup V\&C}$$

(C-SIM)

$$\frac{W_1 \cong W_1' \qquad W_1' \Rrightarrow_{\text{C-}\gamma} W_2' \qquad W_2' \cong W_2 \qquad \gamma \in \{\text{INTER}, \text{MOD}, \text{DEP}\}}{W_1 \Rrightarrow_{\text{C-}\gamma} W_2}$$

Figure 16. Simplification rules for n-ary constrained flux networks, with $\mathcal{I}$ the set of intermediate species and $\mathbf{r}$ the n-tuple of initial reactions.

The rule (C-MOD) removes an intermediate that is never a reactant or a product of a reaction, and replaces X with its initial value $X(0)$. Then the rule (C-DEP) removes a dependent reaction. In contrast to the case without kinetics, the kinetic expressions of the remaining reactions need to be modified. The last rule (C-SIM) states that simplification is applied modulo similarity of constraint flux reaction networks.

The simplification defined here is sound for the contextual equivalence relation of constrained flux networks:

Proposition 2. *Given a constrained flux network W, if $W \Rrightarrow_{\text{C}} W'$ then $W \sim W'$.*

The proof is given in Appendix A. The arguments are direct from the definitions, except that the Diamond Lemma 3 is needed in for $\Rrightarrow_{\text{C-INTER}}$.

7.3. Michaelis-Menten

We illustrate the simplification on the classical Michaelis-Menten example [10].

We consider the simplification of a three-step enzymatic scheme with mass-action kinetics into a single reaction with Michaelis-Menten kinetics. In the initial network, *MMnet* depicted in Figure 17, a substrate S can bind to an enzyme E and form a complex C. The complex can either dissociate back to S and E, or produce a product P, while releasing E. We assume here that the enzyme E and the complex C are intermediate species, i.e., they are at steady-state and cannot interact with the context. Therefore, the intermediate species E and C are linearly steady in this network.

We first look at the elimination of the intermediate C with (C-INTER). To this end, we merge each reaction that produces C (that is, reaction r_1) with each reaction that consumes C (reactions r_2 and r_3) and obtain the network $MMnet_C$. Thus, merging reactions r_1 and r_2 (resp. r_1 and r_3) of *MMnet* (Figure 17) results in the reaction r_{12} (resp. r_{13}) of $MMnet_C$. The simplification also replaces the atomic constraint $cst(C)$ with $cst(k_1SE/(k_2 + k_3))$. Since we also have the constraint S, and the parameters are constant too, we can rewrite $cst(k_1SE/(k_2 + k_3))$ into the similar $cst(S)$. We also add the constraint $C(0) = k_1SE/(k_2 + k_3)$.

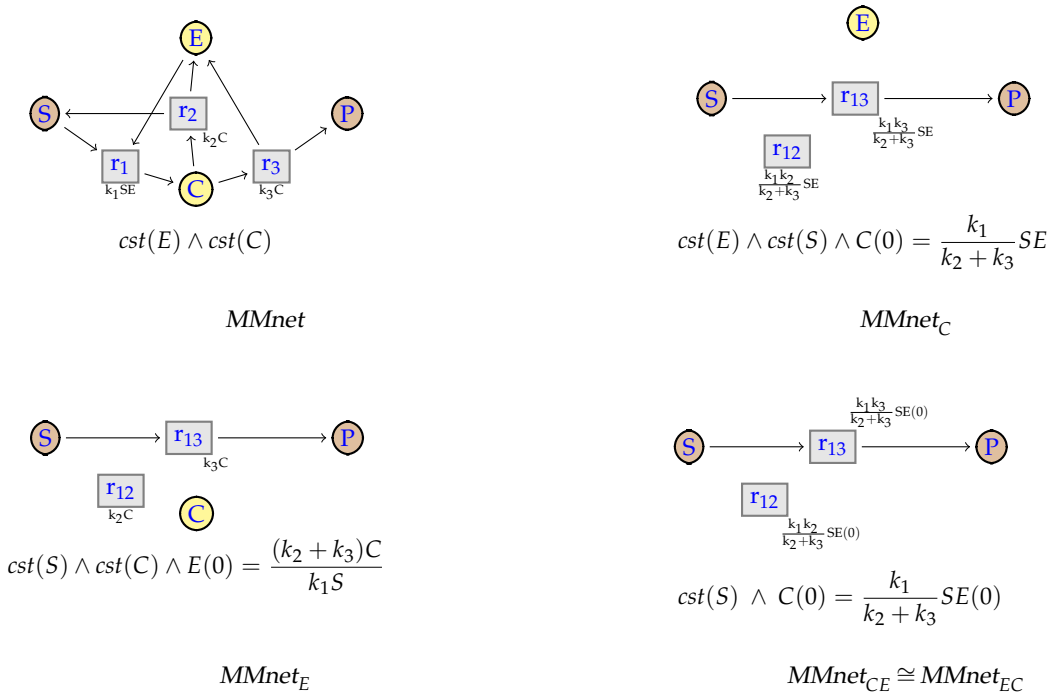


Figure 17. Reaction networks for the Michaelis-Menten example. $MMnet_E$ and $MMnet_C$ are obtained from the initial network *MMnet* after removing E and C respectively. $MMnet_{CE}$ is obtained after removing both C and then E in this order. $MMnet_{EC}$ is obtained by inverting the order of elimination.

To remove E before the elimination of C, one would merge r_3 with r_1, r_2 with r_1, and obtain the network $MMnet_E$. At this point, in both networks we have an intermediate species that is neither a product nor a reactant of any reaction, but is used as a modifier. We can then remove it with (C-Mod), replacing E with $E(0)$ (resp. C with $C(0)$). We obtain respectively the networks $MMnet_{CE}$ and $MMnet_{EC}$. Note that these networks are similar. We can rewrite $C(0) = k_1SE(0)/(k_2+k_3)$ into $E(0) = (k_2+k_3)C(0)/(k_1S)$, and use this equation to rewrite the kinetic rate. Additionally, we can also use it to transform the kinetic expression of r_{12} into the usual one for Michaelis-Menten, using the following transformation. First, we have:

$$\begin{aligned} E(0) &= (E(0)+C(0)) - C(0) \\ &= (E(0)+C(0)) - k_1SE(0)/(k_2+k_3). \end{aligned}$$

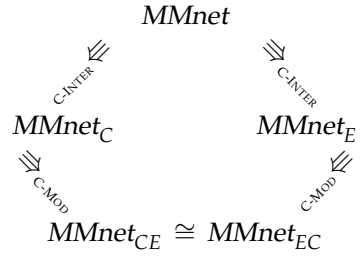
We can then rewrite it into:

$$E(0) = \frac{(E(0)+C(0))(k_2+k_3)}{k_2+k_3+k_1S}.$$

By replacing $E(0)$ with this expression in the kinetic expressions of r_{13} in $MMnet_{CE}$, we obtain after basic rewriting the classical rate:

$$k_3(E(0)+C(0))\frac{S}{\frac{k_2+k_3}{k_1}+S}.$$

The following diagram illustrates the confluence of the simplifications on these networks.



8. Preservation of Linear Steadiness

As we have seen in the previous section, to remove an intermediate species X, we need to impose that X is linearly steady. When removing a set of intermediate species $\mathcal{I}$, we then need that any $X \in \mathcal{I}$ is linearly steady, and moreover than when we remove one intermediate species, the other species in $\mathcal{I}$ remain linearly steady. We therefore introduce the following additional conditions, and denote by *LinNets* the set of networks that satisfy these conditions. We then prove that the set *LinNets* is stable under the simplification, i.e., that the simplification of a network in *LinNets* is still a network in *LinNets*.

8.1. LinNets

We first define the new conditions, and present some examples to motivate them. For any flux $v; ex$, we note $Cons_{\mathcal{I}}(\mathbf{r}_v) = \{X \in \mathcal{I} \mid \mathbf{r}_v(X) < 0\}$ and $Prod_{\mathcal{I}}(\mathbf{r}_v) = \{X \in \mathcal{I} \mid \mathbf{r}_v(X) > 0\}$.

Definition 14. *We denote by LinNets the set of constrained flux networks W such that W is similar to a constrained flux network $V\&C$ and that for all intermediates $X \in \mathcal{I}$:*

1. *Either X is linearly steady in $V\&C$, or X is only a modifier, that is for any $v; e \in V$ we have $\mathbf{r}_v(X) = 0$.*
2. *No intermediate species different from X occurs in the kinetic expression of a reaction that consumes X: for any $v; e \in V$, if $X \in Cons_{\mathcal{I}}(\mathbf{r}_v)$, then $Specs(e) \cap \mathcal{I} \subseteq \{X\}$.*
3. *The rate of a reaction that produces X but does not consume an intermediate species does not depend on the concentration of any intermediate species: for any $v; e \in V$, if $X \in Prod_{\mathcal{I}}(\mathbf{r}_v)$ and $Cons_{\mathcal{I}}(\mathbf{r}_v) = \emptyset$, then $\mathcal{I} \cap Specs(e) = \emptyset$.*
4. *The total stoichiometry of the intermediate species in the reactant (resp. product) of a reaction is never greater than one: for any $v; e \in V$, $|Cons_{\mathcal{I}}(\mathbf{r}_v)| \leq 1$ and $|Prod_{\mathcal{I}}(\mathbf{r}_v)| \leq 1$.*

Note that, as a consequence of the stoichiometry condition, the sets $Cons_{\mathcal{I}}(\mathbf{r}_v)$ and $Prod_{\mathcal{I}}(\mathbf{r}_v)$ are either empty or consist of a single intermediate species.

We illustrate the motivations for these new conditions on the following examples.

Let us first consider the case where a reaction consuming X has a kinetic rate that depends on another intermediate (here Y in the kinetic rate of r_2), so that Condition 2 is not satisfied. It is illustrated in Figure 18.

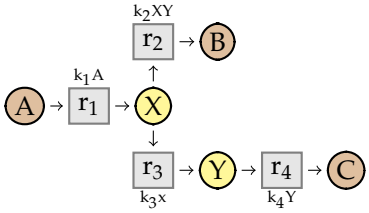


Figure 18. Example illustrating the need of Condition 2.

If we remove Y first by merging r_3 and r_4, then we compute the expression $Y = \frac{k_3}{k_4}X$. We replace Y with this expression in the kinetic rate of r_2, obtaining a reaction with a non-linear kinetic expression $\frac{k_2 k_3}{k_4}X^2$.

Similarly, consider a reaction producing X with a kinetic rate that depends on Y, i.e., a network where Condition 3 is not satisfied (Figure 19).

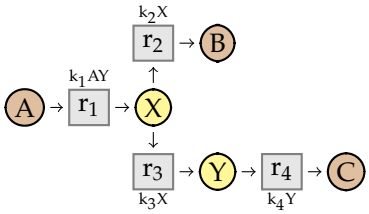


Figure 19. Example illustrating the need of Condition 3.

If we remove the intermediate Y, the kinetic expression of r_1 becomes $\frac{k_1 k_3}{k_4}AX$. We obtain a reaction producing X, with a kinetic expression depending on X. Therefore, the differential equation for X will not have the required form: $0 = \dot{X} = X(\frac{k_1 k_3}{k_4}A - (k_2 + k_3))$, and we cannot compute an expression for X.

This kind of situation may also appear as the result of the simplification of reactions where one intermediate has a stoichiometry greater than one, i.e., Condition 4 is not satisfied (Figure 20).

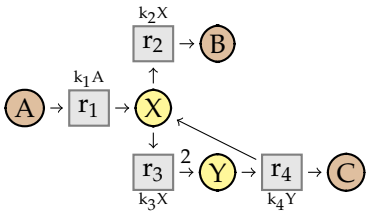


Figure 20. Example illustrating the need of Condition 4.

In this network, reaction r_3 produces two molecules of Y. If we remove Y, the merging of r_3 and r_4 is a reaction that produces one molecule of X (and one of C), with kinetic expression $k_3 X$.

Finally, if we have two intermediate species that are both reactants (or both products) in the same reactions (Condition 4 again), then the stoichiometry of one intermediate can become greater than one as a result of the elimination of the other intermediate (Figure 21).

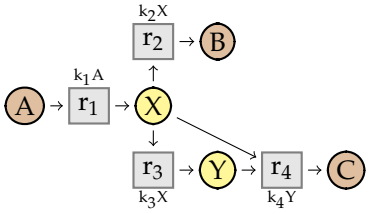


Figure 21. Example illustrating the need of Condition 4.

If we remove Y, the merging of r_3 and r_4 is a reaction with two molecules of X as reactants.

8.2. Stability of LinNets

Now, we prove that *LinNets* is stable for our simplification.

We first consider the following proposition.

Proposition 3. *Let W, W_0 be reaction networks such that $W_0 \in$ LinNets and $W_0 \Rrightarrow_c^* W$. Let $v; e \in W$ be a flux that depends on $v_1; e_1, \ldots, v_k; e_k \in W$. Then:*

- *there exists an index i such that $Cons_{\mathcal{I}}(\boldsymbol{r}_{v_i}) = Cons_{\mathcal{I}}(\boldsymbol{r}_v)$, $Prod_{\mathcal{I}}(\boldsymbol{r}_{v_i}) = Prod_{\mathcal{I}}(\boldsymbol{r}_v)$, and*
- *for any $j \neq i$, $Prod_{\mathcal{I}}(\boldsymbol{r}_{v_j}) = Cons_{\mathcal{I}}(\boldsymbol{r}_{v_j}) = \emptyset$.*

The proof of this proposition is quite long and requires some new notions and definitions, and is given in Appendix B.

We now prove that *LinNets* is stable for the simplification.

Proposition 4. *The set of networks LinNets is stable for the simplification, that is if $W \in$ LinNets and $W \Rrightarrow W'$, then $W' \in$ LinNets.*

Proof. If the simplification is done with (C-MOD) or (C-SIM), then the conditions of *LinNets* are trivially preserved.

Let us assume that the simplification is done with the rule (C-DEP), removing a flux $v; e$ that depends on $v_i; e_i$ with coefficient a_i. So the simplified network contains the fluxes $v_i; e_i + a_i e$. By Proposition 3, there is an i such that $Cons_{\mathcal{I}}(\mathbf{r}_{v_i}) = Cons_{\mathcal{I}}(\mathbf{r}_v)$ and $Prod_{\mathcal{I}}(\mathbf{r}_{v_i}) = Prod_{\mathcal{I}}(\mathbf{r}_v)$, and for any other $j \neq i$, $Cons_{\mathcal{I}}(\mathbf{r}_{v_j}) = Prod_{\mathcal{I}}(\mathbf{r}_{v_j}) = \emptyset$.

The fourth condition on the stoichiometry is trivially preserved by the simplification.

For $j \neq i$, $Cons_{\mathcal{I}}(\mathbf{r}_{v_j}) = Prod_{\mathcal{I}}(\mathbf{r}_{v_j}) = \emptyset$ implies that the conditions on the kinetic expressions are directly satisfied. If $v_i; e_i + a_i e$ consumes X, then since $Cons_{\mathcal{I}}(\mathbf{r}_{v_i}) = Cons_{\mathcal{I}}(\mathbf{r}_v)$, the flux $v; e$ consumes X too. Then by induction, e and e_i are linear in X, and no other intermediate species occurs in them. Therefore, this is also the case for $e_i + a_i e$.

If $v_i; e_i + a_i e$ produces X without consuming any other intermediate, then it is also the case for $v; e$. Then by induction, e, e_i, and $e_i + a_i e$ do not depend on the concentration of any intermediate species. Therefore, the kinetic conditions are satisfied, and $W' \in$ *LinNets*.

Finally, consider the case of a rule (C-INTER) applied on a species X. We denote by *prod* and *cons* the expressions defined in the rule. Since $W \in$ *LinNets*, note that for any $Z \in \mathcal{I}$, we have $Z \notin Vars(cons)$.

Let $v; e \in W'$ be a reaction such that $Cons_{\mathcal{I}}(\mathbf{r}_v) = \{Y\}$. We consider the second condition, on the linearity of Y in e. If $v; e \in W$ (that is the flux has not been changed at all by the simplification rule), then the linearity condition is trivially preserved by induction. $v; e$ cannot be the simplification of a flux $v; e'$ with X as modifier, since that would contradict the linearity condition in W. So now assume the $v; e$ is the merging of a flux $v_p; e_p$ and $v_c; e_c \in W$. Then we have $Cons_{\mathcal{I}}(\mathbf{r}_{v_p}) = \{Y\}$ and $Prod_{\mathcal{I}}(\mathbf{r}_{v_p}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_c}) = \{X\}$. Therefore, the linearity condition implies $e_p = Ye'_p$ and $e_c = Xe'_c$, with for any $Z \in \mathcal{I}$, $Z \notin Spec(e'_p), Spec(e'_c)$. Then we have $e = Ye'_p e'_c / cons$, and the linearity condition is satisfied in W'.

Now let $v; e \in W'$ be a reaction such that $Cons_{\mathcal{I}}(\mathbf{r}_v) = \emptyset$ and $Prod_{\mathcal{I}}(\mathbf{r}_v) = \{Y\}$, and consider the third linearity condition. By linearity, $v; e$ cannot be the simplification of a reaction of W with X as modifier. If we had $v; e \in W$, then the condition is satisfied in W by induction, and therefore in W' too. Assume that $v; e$ is the merging of a reaction $v_p; e_p$ and $v_c; e_c \in W$. Then we have $Cons_{\mathcal{I}}(\mathbf{r}_{v_p}) = \emptyset$, $Prod_{\mathcal{I}}(\mathbf{r}_{v_p}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_c}) = \{X\}$, and $Prod_{\mathcal{I}}(\mathbf{r}_{v_c}) = \{Y\}$. Therefore, the linearity conditions on W imply $e_c = Xe'_c$, with for any $Z \in \mathcal{I}$, $Z \notin Spec(e_p), Spec(e'_c)$. Then we have $e = e_p e_c / cons$, and the condition is satisfied in W'.

Finally, consider the stoichiometric condition. We only have to verify this property for new fluxes $v;e$ that are the merging of a flux $v_p;e_p$ and $v_c;e_c$. We have $Prod_{\mathcal{I}}(\mathbf{r}_{v_p}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_c}) = \{X\}$. Moreover, by normalization, we have

$$Prod_{\mathcal{I}}(\mathbf{r}_{v_p \diamond v_c}) \quad = \quad Prod_{\mathcal{I}}(\mathbf{r}_{v_c}) \backslash Cons_{\mathcal{I}}(\mathbf{r}_{v_p}).$$

Since $|Prod_{\mathcal{I}}(\mathbf{r}_{v_c})| \leq 1$, we have $|Prod_{\mathcal{I}}(\mathbf{r}_{v_p \diamond v_c})| \leq 1$, and similarly for $Cons_{\mathcal{I}}(\mathbf{r}_{v_p \diamond v_c})$. □

Then the set *LinNets* is stable for the simplification. This directly implies that the simplification can remove every intermediate species in $\mathcal{I}$.

9. Confluence of the Simplification Relation

We now study the confluence of the simplification relation. We first show that the structural confluence, that is the confluence of the fluxes without kinetics, is a direct consequence of the previous results. We next present an example that illustrates that, however, the distribution of the kinetics between the fluxes can be different. Finally, we give a criterion on the modes of a network, that guarantees the full confluence, that is confluence of the structure and the rates.

In the following, we only consider networks in *LinNets*.

9.1. Structural Confluence

We say that two constrained flux networks $W = V\&C$ and $W' = V'\&C'$ are *structurally similar*, denoted $W \cong^{struc} W'$, if they have the same structure, that is the same fluxes when neglecting the kinetic expressions:

$$\{v \mid \exists e.\, v;e \in V\} = \{v' \mid \exists e'.\, v';e' \in V'\}.$$

Theorem 3 (Structural confluence)**.** *The relation* $\Rrightarrow_c$ *on* $(LinNets, \cong^{struc})$ *is confluent.*

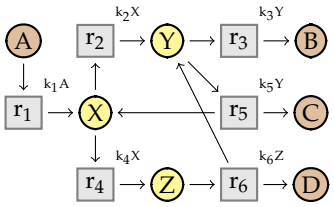
Proof. This is a direct consequence of the stability of the *LinNets* (Proposition 4) and of the confluence of the simplification without kinetics (Theorem 2). □

9.2. Non-Confluence of the Kinetic Rates

Let us consider the reaction network W, depicted on Figure 22, with 7 species and 6 fluxes. The intermediate species are X, Y and Z. Initially, the kinetics are all mass-action.

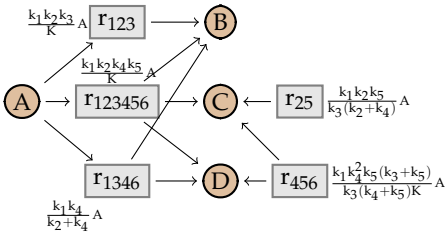
We can remove the intermediate species in different orders with (C-INTER). If we start by eliminating X, followed by Y and finally Z, we obtain the network W_{XYZ}, while if we first eliminate X, then Z and Y, we obtain W_{XZY}. These networks are different, since W_{XYZ} has one additional flux. This illustrates the necessity of the rule (C-DEP) to obtain the same network structure.

Indeed, the additional flux v_{123456} in W_{XYZ} is dependent on v_{123} and v_{456} and can therefore be removed with (C-DEP), while updating the kinetic expressions. We then obtain exactly the network W_{XZY}. However, v_{123456} also depends on v_{25} and v_{1346}. Therefore, we could as well remove it while updating these reactions. In that case, we obtain the different network W_{XYZd}. This network has the same structure as W_{XZY}, but not the same distribution of rates between the fluxes.

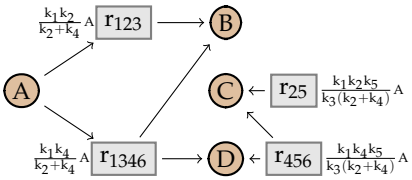


$$cst(X) \wedge cst(Y) \wedge cst(Z)$$

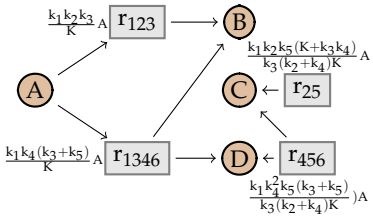
$$W$$



$$cst(A) \wedge X(0) = \frac{k_1(k_3+k_5)}{k_3(k_2+k_4)}A$$
$$\wedge Y(0) = \frac{k_1}{k_3}A \wedge Z(0) = \frac{k_1k_4(k_3+k_5)}{k_3k_6(k_2+k_4)}A$$
$$W_{XYZ}$$



$$cst(A) \wedge X(0) = \frac{k_1(k_3+k_5)}{k_3(k_2+k_4)}A$$
$$\wedge Y(0) = \frac{k_1}{k_3}A \wedge Z(0) = \frac{k_1k_4(k_3+k_5)}{k_3k_6(k_2+k_4)}A$$
$$W_{XZY}$$



$$cst(A) \wedge X(0) = \frac{k_1(k_3+k_5)}{k_3(k_2+k_4)}A$$
$$\wedge Y(0) = \frac{k_1}{k_3}A \wedge Z(0) = \frac{k_1k_4(k_3+k_5)}{k_3k_6(k_2+k_4)}A$$
$$W_{XYZd}$$

Figure 22. Network W and its simplifications. (**top left**) Network W. (**top right**) Network W_{XYZ} after eliminating X, Y and Z (in this order). (**bottom left**) Network W_{XZY} after eliminating X, Z and Y. (**bottom right**) Network W_{XYZd} after eliminating X, Y, Z, and the dependent reaction. The new parameter is $K = k_2k_3 + k_3k_4 + k_4k_5$.

9.3. Criterion for the Full Confluence

We now give a criterion that guaranties the full confluence of the simplification.

Definition 15. *A vector of reactions* $\boldsymbol{r} = (r_1, \ldots, r_n)$ *is* uniquely decomposable *if any mode* $v \in ker_+(S)$ *has an unique decomposition in elementary modes, where* S *is the stoichiometric matrix of* $\mathtt{N}_{\boldsymbol{r}|\mathcal{I}}$.

Example 6. *Consider the network* W *represented in the Figure 22, with* $\mathcal{I} = \{X, Y, Z\}$. *It has 4 different elementary modes:* $v_1 = (1,1,1,0,0,0)$, $v_2 = (1,0,1,1,0,1)$, $v_3 = (0,1,0,0,1,0)$ *and* $v_4 = (0,0,0,1,1,1)$. *Then the mode* $v = (1,1,1,1,1,1)$ *can be decompose in either* $v_1 + v_4$, *or in* $v_2 + v_3$. *The two decompositions are illustrated in Figure 23.* $\boldsymbol{r}$ *is not uniquely decomposable, and the simplification is not confluent for the kinetic rates, as we have seen before.*

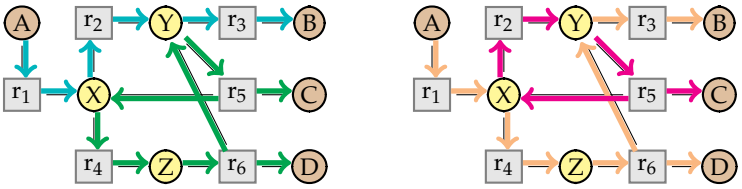


Figure 23. Two decompositions of the mode $(1,1,1,1,1,1)$ in the network W.

Theorem 4 (Confluence). *If the initial vector of reactions* $\mathbf{r}$ *is uniquely decomposable, then the relation* $\Rrightarrow_c$ *on* $(LinNets, \cong)$ *is confluent, for both the structure and the kinetic rates.*

The theorem is the consequence of the following lemmas, that analyze the different critical pairs. That is, if a network W can be simplified in two different manners into W_1 and W_2, then these two networks can be simplified into W_1' and W_2' such that $W_1' \cong W_2'$.

Lemma 7. *Assume* $\mathbf{r}$ *is uniquely decomposable. Let* W *be a network such that* $W \Rrightarrow_{\text{C-Dep}} W_i$ *for* $i \in \{1,2\}$. *Then* $\exists W_i'$ *such that* $W_i \Rrightarrow_c^* W_i'$ *and* $W_1' \cong W_2'$.

Proof. Let $v_i e_i$ be the dependent flux removed when simplifying W into W_i, for $i \in \{1,2\}$, with $v_i; e_i$ dependent on $v_i^1; e_i^1, \ldots, v_i^{k_i}; e_i^{k_i}$ with coefficients $a_i^1, \ldots, a_i^{k_i}$.

If $v_1 = v_2$, since $\mathbf{r}$ is uniquely decomposable, we have $\{v_1^1, \ldots, v_1^{k_1}\} = \{v_2^1, \ldots, v_2^{k_2}\}$. So the simplified networks are trivially the same, that is $W_1 = W_2$.

Assume $v_1 \neq v_2$, and that for any $i \in \{1,2\}$, for any j, $v_i \neq v_{3-i}^j$, that is v_1 does not depend on v_2 and reciprocally. Then we can still remove v_1 in W_2, and v_2 in W_1, and we find the same network modulo similarity: $W_1' \cong W_2'$.

If v_1 depends on v_2, and v_2 depends on v_1, then since dependencies are positive linear combinations, that directly implies $v_1 = v_2$.

Finally, if $v_1 \neq v_2$, and v_1 depends on v_2, with coefficient a, but v_2 does not depend on v_1 (or conversely), we have $v_1 = \sum_j a_1^j v_1^j + a v_2$ and $v_2 = \sum_j a_2^j v_2^j$. If we remove v_1, $v_2; e_2$ becomes $v_2; e_2 + a e_1$, and can be removed. The fluxes obtained are of the form $v^j; e^j + a_1^j e_1 + a_2^j (e_2 + a e_1)$. If we remove v_2 first, then we can remark that v_1 is still dependent, with $v_1 = \sum_j a_1^j v_1^j + a(\sum_j a_2^j v_2^j)$, and can be removed. We obtain the fluxes $v^j; e^j + a_2^j e_2 + (a_1^j + a a_2^j) e_1$. Then the simplified fluxes are similar, and $W_1' \cong W_2'$. □

Lemma 8. *Let* W *be a network such that* $W \Rrightarrow_{\text{C-Dep}} W_1$ *and* $W \Rrightarrow_{\text{C-Mod}} W_2$. *Then* $\exists W_i'$ *such that* $W_i \Rrightarrow_c^* W_i'$ *and* $W_1' \cong W_2'$.

Proof. This case is quite trivial. Let $v; e$ be the dependent flux, X the modifier, and $v'; e'$ another reaction such that v depends on v' with factor a. If we remove X first and then v, the flux $v'; e'$ is simplified into $v'; e'[X := X(0)] + a e[X := X(0)]$. Otherwise, it is simplified into $v'; (e' + a e)[X := X(0)]$. The two expressions are trivially similar. □

Lemma 9. *Let* W *be a network such that* $W \Rrightarrow_{\text{C-Dep}} W_1$ *and* $W \Rrightarrow_{\text{C-Inter}} W_2$. *Then* $\exists W_i'$ *such that* $W_i \Rrightarrow_c^* W_i'$ *and* $W_1' \cong W_2'$.

Proof. The full proof is given in Appendix C. The main idea is to use Proposition 3 to prove that if X is in the dependent flux v_d, it is also in one of the fluxes v_i that v_d depends on. Therefore, if we eliminate X and combine v_d with another flux v', we also merge v_i with v'. Then $v_d \diamond_X v'$ is still dependent on $v_i \diamond_X v'$ and other fluxes, and can be removed. □

Lemma 10. *Let W be a network such that $W \Rightarrow_{\text{C-MOD}} W_i$ for $i \in \{1,2\}$. Then $\exists W_i'$ such that $W_i \Rightarrow_C^* W_i'$ and $W_1' \cong W_2'$.*

Proof. This case is trivial, since the substitutions commute. □

Lemma 11. *Let W be a network such that $W \Rightarrow_{\text{C-MOD}} W_1$ and $W \Rightarrow_{\text{C-INTER}} W_2$. Then $\exists W_i'$ such that $W_i \Rightarrow_C^* W_i'$ and $W_1' \cong W_2'$.*

Proof. Once again, since the two removed species cannot be the same, this case is trivial. □

Lemma 12. *Let W be a network such that $W \Rightarrow_{\text{C-INTER}} W_i$ for $i \in \{1,2\}$. Then $\exists W_i'$ such that $W_i \Rightarrow_C^* W_i'$ and $W_1' \cong W_2'$.*

Proof. The full proof is given in Appendix C. The idea is that after removing one intermediate species, we can still remove the other one, either with (C-MOD) or with (C-INTER). In the second case, some dependent fluxes are generated, that we can eliminate to find the same simplified network, whatever the order of elimination of the intermediate species. □

10. An Example from the BioModels Database

We have shown that the simplification system that we presented can exhibit non-confluence of the rates, even in a simple scenario with a small number of intermediates. To find if such a situation occurs in practice, we investigated the SBML models in the curated BioModels database [6]. We were thus able to find a network, for the model BIOMD0000000173, that does not verify the confluence criterion, and such that two different simplified networks can be identified. Note that this was the only model not satisfying the criterion, when considering every model of the BioModels database with mass-action kinetics and with three or four linear intermediate species.

The network identified is a model of the *Smad*-based signal transduction mechanisms from the cell membrane to the nucleus, presented in [23]. We only consider here a sub-network W of this model, sufficient to illustrate the non-confluence. It is represented in Figure 24.

In this network, a molecule of $S4_c$, that represents the species *Smad*4 in the cytoplasm, can bind with either a molecule of *Smad*2 in a phosphorylated form ($pS2_c$) and form the complex $S24_c$ (reaction r_5), or with a molecule of G in a phosphorylated form (pG_c), and form the complex $G4_c$ (reaction r_{22}). These two reactions are reversible ($r_{5'}$ and $r_{22'}$). The same transformations can occur in the nucleus (reactions r_6, $r_{6'}$, r_{23} and $r_{23'}$). The species *Smad*4 can also move from the cytoplasm to the nucleus, or reciprocally (r_1 and $r_{1'}$). Finally, the complex of *Smad*2 and *Smad*4 can move from the cytoplasm to the nucleus (r_7).

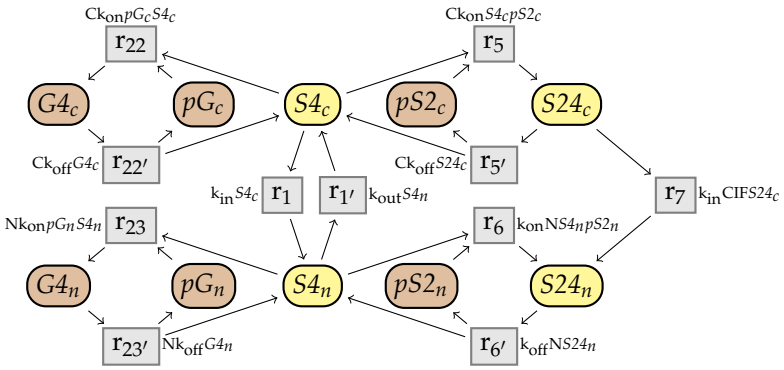
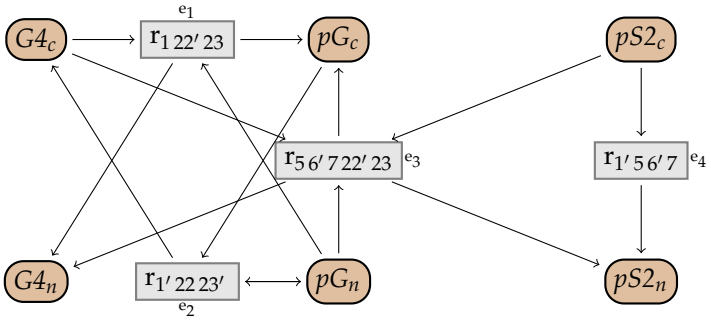


Figure 24. Sub-network W of the *Smad*-based signal transduction model from [23].

We assume that $\mathcal{I} = \{S4_c, S24_c, S4_n, S24_n\}$. The network is in *LinNets*. Therefore, we can consider the elimination of the four intermediate species. According to the order of the simplification, we can then obtain two different networks, with the same structure, but with different kinetic expressions. They are represented in Figure 25. The network W_1 is obtained by removing $S4_n$ first, then $S24_n$, then $S24_c$, and finally $S4_c$ and the dependent fluxes. The network W_2 is obtained by removing $S4_n$, then $S24_n$, $S4_c$, $S24_c$ and the dependent fluxes.



	In W_1	In W_2
e_1	$\frac{2G4_cpG_n}{K_1}$	$\frac{G4_cpG_n(K_1 + K_3pS2_c)}{K_1K_2}$
e_2	$\frac{2G4_npG_c}{K_1}$	$\frac{2G4_npG_c}{K_1}$
e_3	$\frac{G4_cpG_npS2_c}{K_1}$	$\frac{G4_cpG_npS2_c(pG_c + pS2_c + 1)K_3}{K_1K_2}$
e_4	$\frac{pS2_c(G4_c + G4_n)}{K_1}$	$\frac{pS2_c(G4_c(pS2_c + pG_c)K_3 + G4_nK_2)}{K_1K_2}$

Parameters : $K_1 = 2G4_c + 2pG_c + 2G4_cpG_c + pG_cpS2_c$
$K_2 = G4_c + pG_c + pS2_c + G4_cpG_c + pS2_cpG_c$
$K_3 = 1 + pG_c$

Figure 25. Simplified networks from W. Both networks have the same structure, and the kinetic expressions are defined in the table. The network W_1 is obtained by removing in order $S4_n$, $S24_n$, $S24_c$, $S4_c$ and the dependent fluxes. The network W_2 is obtained by removing in order $S4_n$, $S24_n$, $S4_c$, $S24_c$ and the dependent fluxes.

We now show that the criterion is not satisfied in the initial network, i.e., that W is not uniquely decomposable. We consider the following mode:

$$v = v_{22'} + v_1 + v_{1'} + v_5 + v_7 + v_{6'} + v_{23}.$$

This mode has two possible decompositions into elementary modes, $\{v_{red}, v_{blue}\}$ and $\{v_{green}, v_{magenta}\}$, with:

$$\begin{aligned} v_{red} &= v_{22'} + v_5 + v_7 + v_{6'} + v_{23}, \\ v_{blue} &= v_1 + v_{1'}, \\ v_{green} &= v_{22'} + v_1 + v_{23}, \\ v_{magenta} &= v_{1'} + v_5 + v_7 + v_{6'}. \end{aligned}$$

We represent these fluxes in Figure 26, where we omit the non-intermediate species and the kinetic expressions for the sake of simplicity.

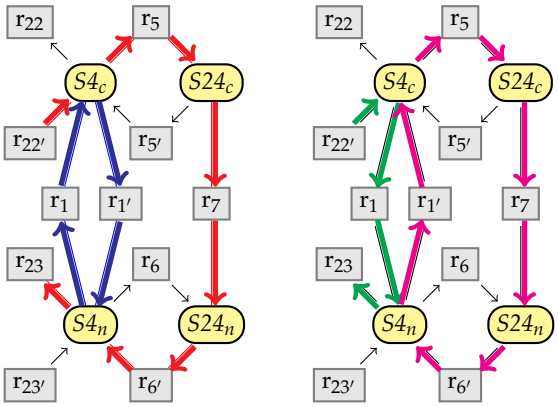


Figure 26. In red the elementary mode v_{red}, in blue v_{blue}, in green v_{green}, and in magenta $v_{magenta}$.

11. Simplification of Systems of Equations

In this section, we study the relation between the simplification $\Rrightarrow_c$ on reaction networks and a simplification $\Rightarrow$ on systems. We show that the assignment of a system $E(W)$ to a network W is a simulation for the simplifications.

11.1. Simplification of Systems of Equations

The simplification of systems is illustrated in Figure 27. The first rule replaces a constant variable x with its initial value $x(0)$, and $\dot{x}$ with 0. The second rule extends the simplification to similar systems. We define the simplification:

$$\Rightarrow \quad =_{\text{df}} \quad \Rightarrow_{\text{E-INTER}} .$$

$$\text{(E-INTER)} \qquad \frac{E \models cst(x)}{E \Rightarrow_{\text{E-INTER}} E[x := x(0)][\dot{x} := 0]}$$

$$\text{(E-SIM)} \qquad \frac{E_1 \cong E_1' \qquad E_1' \Rightarrow_{\text{E-INTER}} E_2' \qquad E_2' \cong E_2}{E_1 \Rightarrow_{\text{E-INTER}} E_2}$$

Figure 27. Simplification rules for systems of equations.

Lemma 13. *The simplification is correct for the equivalence, that is:*

$$E \Rightarrow E' \qquad \textit{implies} \qquad E \simeq E'.$$

Theorem 5. *The relation $\Rightarrow$ on $(Syst, \cong)$ is uniformly confluent.*

Proof. It is trivial, since the substitutions commute. □

11.2. Simulation

The assignment of a system $E(W)$ to a network W is a simulation from $(LinNets, \Rrightarrow_c)$ to $(Systems, \Rightarrow^*)$.

Lemma 14. *Given a network $W \in LinNets$, if $W \Rrightarrow_c W'$, then $E(W) \Rightarrow^* E(W')$.*

Proof. The rule (C-SIM) for the networks is directly imitated by the rule (E-SIM) for the systems.

For the rule (C-MOD), if $W \Rrightarrow_{\text{C-MOD}} W'$, then we directly have $E(W) \Rightarrow_{\text{E-INTER}} E(W')$.

For the rule (C-INTER), assume that we remove a species X from W to obtain W'. Then $E(W) \models cst(x_X)$, therefore, we can simplify $E(W)$ into a system E'. We have to prove that $E' \sim E(W')$. First, observe that the systems have the same variables.

Consider a differential equation of $E(W)$, for a species $A \neq X$:

$$\begin{aligned}\dot{A} &= \sum_{v;e\in W} \mathbf{r}_v(A)e \\ &= \sum_{\{v;e\in W|\mathbf{r}_v(X)>0\}} \mathbf{r}_v(A)e \quad + \sum_{\{v;e\in W|\mathbf{r}_v(X)<0\}} \mathbf{r}_v(A)e \quad + \sum_{\{v;e\in W|\mathbf{r}_v(X)=0\}} \mathbf{r}_v(A)e.\end{aligned}$$

We define $prod = \sum_{\{v;e\in W|\mathbf{r}_v(X)>0\}} e$ and $cons = \sum_{\{v;e\in W|\mathbf{r}_v(X)<0\}} e$. The differential equation in $E(W')$ becomes:

$$\begin{aligned}\dot{A} &= \sum_{v;e\in W'} \mathbf{r}_v(A)e \\ &= \sum_{\{v;e,v';e'\in W|\mathbf{r}_v(X)>0,\mathbf{r}_{v'}(X)<0\}} \mathbf{r}_{v\diamond_X v'}(A)\frac{ee'}{cons} \quad + \sum_{\{v;e\in W|\mathbf{r}_v(X)=0\}} \mathbf{r}_v(A)e[X := X(0)] \\ &= \sum_{\{v;e,v';e'\in W|\mathbf{r}_v(X)>0,\mathbf{r}_{v'}(X)<0\}} (\mathbf{r}_v(A) + \mathbf{r}_{v'}(A))\frac{ee'}{cons} \quad + \sum_{\{v;e\in W|\mathbf{r}_v(X)=0\}} \mathbf{r}_v(A)e[X := X(0)] \\ &= \sum_{\{v;e\in W|\mathbf{r}_v(X)>0\}} \mathbf{r}_v(A)\frac{e}{cons}\Big(\sum_{\{v';e'\in W|\mathbf{r}_{v'}(X)<0\}} e'\Big) + \sum_{\{v';e'\in N|\mathbf{r}_{v'}(X)<0\}} \mathbf{r}_{v'}(A)\frac{e'}{cons}\Big(\sum_{\{v;e\in W|\mathbf{r}_v(X)>0\}} e\Big) \\ &\quad + \sum_{\{v;e\in W|\mathbf{r}_v(X)=0\}} \mathbf{r}_v(A)e[X := X(0)] \\ &= \sum_{\{v;e\in W|\mathbf{r}_v(X)>0\}} \mathbf{r}_v(A)\frac{econs}{cons} + \sum_{\{v';e'\in W|\mathbf{r}_{v'}(X)<0\}} \mathbf{r}_{v'}(A)\frac{e'prod}{cons}\Big) + \sum_{\{v;e\in W|\mathbf{r}_v(X)=0\}} \mathbf{r}_v(A)e[X := X(0)].\end{aligned}$$

The system $E(W')$ also contains the constraint $X(0) = X\frac{prod}{cons}$. In addition, by the linearity conditions, we know that for any $v;e \in W$ such that $\mathbf{r}_v(X) > 0$, we have $X \notin Vars(e)$, so $e = e[X := X(0)]$. For any $v';e' \in W$ such that $\mathbf{r}_{v'}(X) < 0$, we have $e' = Xe''$, with $X \notin Vars(e'')$. Therefore, $\frac{e'prod}{cons} = e''\frac{Xprod}{cons} = e''X(0) = e'[X := X(0)]$. So we can rewrite the previous differential equation into:

$$\dot{A} = \sum_{v;e\in W} \mathbf{r}_v(A)e[X := X(0)].$$

In E', we directly have

$$\dot{A} = \sum_{v;e\in W} \mathbf{r}_v(A)e[X := X(0)].$$

Moreover, in E', the equation $\dot{X} = prod - cons$ is replaced by $0 = prod[X := X(0)] - cons[X := X(0)]$. We then have $prod[X := X(0)] = prod$, while $cons = Xe$ for some e such that $X \notin Vars(e)$. Then $cons[X := X(0)] = Xe[X := X(0)] = X(0)e = \frac{X(0)}{X}cons$. So we can rewrite the equation $0 = prod[X := X(0)] - cons[X := X(0)]$ into $0 = prod - \frac{X(0)}{X}cons$, and then into $X(0) = \frac{Xprod}{cons}$. Therefore, the two systems E' and $E(W')$ have the same differential equations and the same constraint, and they are similar.

Finally, consider the rule (C-DEP). Let $v;e$ be the removed reaction, depending on $v_1;e_1,\ldots,v_k;e_k$, with coefficients $a_1,\ldots,a_k$. We write V' for the set of the other fluxes in W. Let A be a species. The ordinary differential equation for A in $E(W)$ is:

$$\dot{A} = \mathbf{r}_v(A)e + \sum_{1\leq i\leq k} \mathbf{r}_{v_i}(A)e_i + \sum_{v';e'\in V'} \mathbf{r}_{v'}(A)e'.$$

Since we have $\mathbf{r}_v = \sum_{1\leq i\leq k} a_i \mathbf{r}_{v_i}$, the equation is similar to:

$$\begin{aligned}\dot{A} &= \sum_{1\leq i\leq k} a_i \mathbf{r}_{v_i}(A)e + \sum_{1\leq i\leq k} \mathbf{r}_{v_i}(A)e_i + \sum_{v';e'\in V'} \mathbf{r}_{v'}(A)e' \\ &= \sum_{1\leq i\leq k} \mathbf{r}_{v_i}(A)(e_i + a_i e) + \sum_{v';e'\in V'} \mathbf{r}_{v'}(A)e'.\end{aligned}$$

This is the equation for A in the system $E(W')$ for the simplified network. Therefore, $E(W) \cong E(W')$. □

12. Conclusions

We have first shown that when neglecting the kinetic expressions, the elimination of linear intermediate species and dependent reactions is a reformulation of the double description method, that computes the elementary modes, and therefore that the network structure of simplified networks is unique. In a second time, when considering kinetic expressions, we provided a biological example illustrating that the simplification can produce two networks with the same structure but different kinetics. We then gave a sufficient criterion on the network structure of the initial network that guarantees the confluence of both the structure and the rates.

Note that the criterion seems to be satisfied in most cases in practice. When looking at the networks with mass-action kinetics from the BioModels database [6], and considering at most four intermediate species, only the *Smad*-model BIOMD0000000173 was identified as not satisfying the criterion. On the other hand, the linearly steadiness as well as the conditions required for a network to be in *LinNets* (such as the stoichiometry conditions, etc.) are not always satisfied in real biological networks, and these are therefore a real restriction on our simplification approach.

Acknowledgments: This work has been funded by the French National Research Agency research grant Iceberg ANR-IABI-3096.

Author Contributions: All authors contributed equally to this work and to the writing of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Soundness of the Simplification Rules for Constrained Flux Networks

We prove Proposition 2, stating that the simplification is sound for the congruence: $W \Rrightarrow_{\mathrm{C}} W'$ implies $W \sim W'$.

We prove, for each simplification rule, that for any context W'' with $Spec(W'') \cap \mathcal{I} = \emptyset$ and for any assignment α, we have $\alpha \in sol(E(W \mid W''))$ iff $\alpha \in sol(E(W' \mid W''))$. Let us assume that $W = V\&C$, $W' = V'\&C'$, and $W'' = V''\&C''$. We therefore have

$$E(W \mid W'') = (\bigwedge_{A\in Spec} \dot{A} = e_A^V + e_A^{V''}) \ \wedge \ C \wedge C''$$

where

$$e_A^V = \sum_{v;e\in V} \mathbf{r}_v(A)e \text{ and } e_A^{V''} = \sum_{v;e\in V''} \mathbf{r}_v(A)e.$$

(C-Mod) Suppose that the removed species is $X \in \mathcal{I}$. Since $Spec(W'') \cap \mathcal{I} = \emptyset$ for any $A \in Spec$, $X \notin Vars(e_A^{V''}) \cup Vars(C'')$ and $e_X^{V''} = 0$. Moreover, $\forall v; e \in V.\mathbf{r}_{\mathbf{r}_v}(X) = 0$, thus $e_A^V = 0$ and the ODE for X in $E(W \mid W'')$ is $\dot{X} = 0$, which is also the equation for X in $E(W' \mid W'')$. As a consequence any solution α should verify $X = X(0)$. In addition, since $E(W \mid W'')$ and $E(W' \mid W'')$ only differ by the substitution of $X(0)$ for X, they have the same solutions.

(C-DEP) Let $V = V_0 \cup \{v; e\} \cup \{v_i; e_i \mid 1 \leq i \leq n\}$ where $v = \sum_{1 \leq i \leq n} a_i v_i$, i.e., v depends on $v_1, \ldots, v_n$ with coefficients $a_1, \ldots, a_n$. For any species A, we can write its ODE in $E(W \mid W'')$ as

$$\begin{aligned}
\dot{A} &= \sum_{v'; e' \in V_0 \cup V''} \mathbf{r}_{v'}(A) e' + \mathbf{r}_v(A) e + \sum_{1 \leq i \leq n} \mathbf{r}_{v_i}(A) e_i \\
&= \sum_{v'; e' \in V_0 \cup V''} \mathbf{r}_{v'}(A) e' + (\sum_{1 \leq i \leq n} a_i \mathbf{r}_{v_i}(A)) e + \sum_{1 \leq i \leq n} \mathbf{r}_{v_i}(A) e_i \\
&= \sum_{v'; e' \in V_0 \cup V''} \mathbf{r}_{v'}(A) e' + \sum_{1 \leq i \leq n} \mathbf{r}_{v_i}(A)(a_i e_i + e)
\end{aligned}$$

which is exactly the ODE for A in $E(W' \mid W'')$. In addition, the constraints in $E(W \mid W'')$ and $E(W' \mid W'')$ are the same. Therefore, their solutions are identical.

(C-SIM) The soundness of this rule comes directly from Lemma 6.

(C-INTER) Suppose that the removed species is $X \in \mathcal{I}$. We note that $e_X^{V''} = 0$.

Let $(v_1, \ldots, v_k)$ and $(e_1, \ldots, e_k)$ be such that $V = \{v_1; e_1, \ldots, v_k; e_k)$. Let P, C, *prod*, and *cons* be as in the Diamond Lemma 3, with $\mathcal{G}$ the set of kinetic expressions, $g_i = e_i$ for all $1 \leq i \leq k$, and $h : \mathbb{N}^n \to \mathcal{G}$ the homomorphism with $h(v) = \mathbf{r}_v(A)$. P is the set of indices for fluxes that produce X, C the set of indices for fluxes that consume X, while *prod* is the total rate of production of X, and *cons* its total rate of consumption.

The ODE for X in $E(W \mid W'')$ is $\dot{X} = prod - cons$. Since X is linearly steady in W, it follows that $C \models (prod = cons) \wedge cons \neq 0$. Therefore, we have that $C \models\!\models C[X := X(0)] \wedge X(0) = X^{\frac{prod}{cons}}$. Let $A \neq X$ be a species. The ODE for A in $E(W \mid W'')$ writes as

$$\dot{A} = \sum_{p \in P} \mathbf{r}_{v_p}(A) e_p + \sum_{c \in C} \mathbf{r}_{v_c}(A) e_c + \sum_{\{v; e \in V \cup V'' \mid \mathbf{r}_v(X) = 0\}} \mathbf{r}_v(A) e.$$

We next consider the ODE for A in $E(W' \mid W'')$ and show that it can be rewritten to obtain the same result. We have

$$\begin{aligned}
\dot{A} &= \sum_{p \in P} \sum_{c \in C} \mathbf{r}_{v_p \diamond_X v_c}(A) \frac{e_p e_c}{cons} + \sum_{m \in M} \mathbf{r}_{v_m}(A) e_m [X := X(0)] + \sum_{v; e \in V''} \mathbf{r}_v(A) e \\
&= \tfrac{1}{cons} \sum_{p \in P} \sum_{c \in C} e_p e_c h(v_p \diamond_X v_c) + \sum_{\{v; e \in V \cup V'' \mid \mathbf{r}_v(X) = 0\}} \mathbf{r}_v(A) e [X := X(0)] \qquad (X \notin pSpecs(V'') \text{ by compatibility}) \\
&= \tfrac{1}{cons} (\sum_{p \in P} e_p \; cons \; h(v_p) + \sum_{c \in C} e_c \; prod \; h(v_c)) + \sum_{\{v; e \in V \cup V'' \mid \mathbf{r}_v(X) = 0\}} \mathbf{r}_v(A) e [X := X(0)] \qquad \text{(Diamond Lemma 3)} \\
&= \sum_{p \in P} e_p \mathbf{r}_{v_p}(A) + \sum_{c \in C} e_c \mathbf{r}_{v_c}(A) + \sum_{\{v; e \in V \cup V'' \mid \mathbf{r}_v(X) = 0\}} \mathbf{r}_v(A) e [X := X(0)] \qquad (C \models prod = cons \wedge cons \neq 0).
\end{aligned}$$

Without the substitution $[X := X(0)]$, this is indeed the equation of A in $E(W \mid W'')$. The substitution is permitted since we argue modulo similarity of constrained flux networks, and since $C \models cst(X)$. Therefore, $E(W \mid W'') \models\!\models E(W' \mid W'')$ so that $W \sim W'$.

Appendix B. Proofs for the Stability of *LinNets*

We first need to introduce some new notions.

We write $\mathbf{r} = (r_1, \ldots, r_n)$ for the initial vector of reactions, and $(v_1, \ldots, v_n)$ for the corresponding unit vectors, that is, for any i, $\mathbf{r}_{v_i} = r_i$. Let $W_0 \in$ *LinNets* be a constrained flux network, and W a network obtained by simplifying W_0, that is $W_0 \Rrightarrow_c^* W$.

We first introduce the notion of paths. We will then relate them to the fluxes in W. Note that we need to distinguish between the case of circular and the case of non-circular path.

Definition A1. *Let* $\mathbf{r}$ *be the initial vector of reactions.*

- *a* path $\tilde{v} = v_1 \ldots v_k$ *is a (non empty) sequence of unit vectors* $v_i \in \mathbb{N}^n$*, such that for any* $1 \leq i < k$*, we have* $Prod_{\mathcal{I}}(\mathbf{r}_{v_i}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_{i+1}}) \neq \emptyset$*;*

- *we denote the vector of a path by* $\sum \tilde{v} = \sum_{1 \leq i \leq k} v_i$;
- *a path is* circular *if* $Prod_{\mathcal{I}}(\boldsymbol{r}_{v_k}) = Cons_{\mathcal{I}}(\boldsymbol{r}_{v_1}) \neq \emptyset$, *and* non-circular *otherwise;*
- *for a circular path* $\tilde{v}$, *we denote the number of intermediate species occurring in the path by* $\tilde{v}(X) = |\{1 \leq i \leq k \mid Prod_{\mathcal{I}}(\boldsymbol{r}_{v_i}) = \{X\}\}|$;
- *for a non-circular path* $\tilde{v}$, *we denote its beginning and its end by* $Cons_{\mathcal{I}}(\boldsymbol{r}_{\tilde{v}}) = Cons_{\mathcal{I}}(\boldsymbol{r}_{v_1})$ *and* $Prod_{\mathcal{I}}(\boldsymbol{r}_{\tilde{v}}) = Prod_{\mathcal{I}}(\boldsymbol{r}_{v_k})$. *In addition, we define the multiset* $\tilde{v}(X) = |\{1 \leq i < k \mid Prod_{\mathcal{I}}(\boldsymbol{r}_{v_i}) = Cons_{\mathcal{I}}(\boldsymbol{r}_{v_{i+1}}) = \{X\}\}|$. *Note that we do not count* $Cons_{\mathcal{I}}(\boldsymbol{r}_{\tilde{v}})$ *and* $Prod_{\mathcal{I}}(\boldsymbol{r}_{\tilde{v}})$ *in this multiset.*

Example A1. *For instance, consider the initial network* W_0, *with* $\mathcal{I} = \{X, Y, Z\}$, *in Figure A1 (left). We denote by* v_i *the unit vector for reaction* r_i.

The path $v_1v_2v_3$ *is non-circular. It has for vector* $(1, 1, 1, 0, 0)$. *We have* $Cons_{\mathcal{I}}(\boldsymbol{r}_{v_1v_2v_3}) = \emptyset$ *and* $Prod_{\mathcal{I}}(\boldsymbol{r}_{v_1v_2v_3}) = Z$. *The multiset is defined by* $v_1v_2v_3(X) = v_1v_2v_3(Y) = 1$, *and* $v_1v_2v_3(Z) = 0$ *(since Z is the end of the path, and not an intermediate node).*

The path v_3v_4 *is circular. It has for vector* $(0, 0, 1, 1, 0)$. *The multiset is defined by* $v_3v_4(X) = 0$ *and* $v_3v_4(Y) = v_3v_4(Z) = 1$.

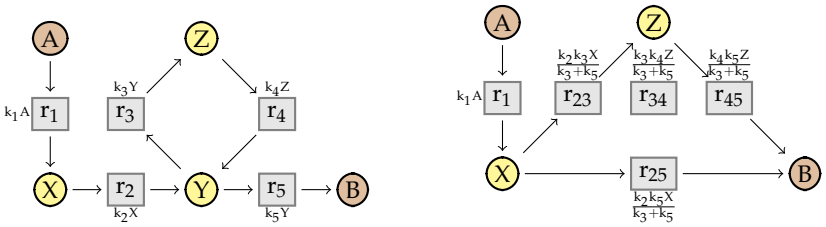


Figure A1. Networks W_0 **(left)**; and W **(right)**.

We will see later that if a path $\tilde{v}$ satisfies some conditions w.r.t. some reaction network W, then there is a corresponding flux v in W such that $\sum \tilde{v} = v$. The reciprocal property also holds. Such a particular path, called a *flux-path*, is formally defined as follows.

Definition A2. *Let* W, W_0 *be reaction networks such that* $W_0 \Rrightarrow_c^* W$,

- *a* non-circular flux-path $\tilde{v}$ *is a non-circular path in* W *such that for any intermediate species* X *with* $\tilde{v}(X) > 0$, *we have* $X \in Spec(W_0) \backslash Spec(W)$ *(meaning that one of the simplification steps* $W_0 \Rrightarrow_c^* W$ *removes* X *from* W_0*), and such that* $Cons_{\mathcal{I}}(\boldsymbol{r}_{\tilde{v}})$ *and* $Prod_{\mathcal{I}}(\boldsymbol{r}_{\tilde{v}})$ *are either the empty solution* $\emptyset$, *or the intermediate species that are still in* W,
- *a* circular flux-path $\tilde{v}$ *is a circular path in* W *if there is at most one intermediate species* X *such that* $\tilde{v}(X) > 0$ *and* $X \in Spec(W)$ *(i.e., X is not yet simplified),*
- *we call* flux-path *a path that is either a circular or a non-circular flux-path.*
- *a flux-path* $\tilde{v}$ *is said to* correspond *to a flux* v *in* W *if* $\sum \tilde{v} = v$.

Example A2. *Let us consider the simplified network* W *in Figure A1 (right).*

The non-circular path $v_1v_2v_3$ *is not a non-circular flux-path, since* $v_1v_2v_3(X) > 0$ *and* X *has not been removed. The path* v_2v_3 *is a non-circular flux-path, and corresponds to the flux* v_{23}.

The path v_3v_4 *is a circular flux-path, since there is a unique intermediate species,* Z, *that has not been removed and such that* $v_3v_4(Z) > 0$. *It corresponds to* v_{34}.

We first prove the following lemma on the dependent flux.

Lemma A1. *Let $v;e$ be a flux that depends on $v_1;e_1,\dots,v_k;e_k$ with coefficients $a_1,\dots,a_k$. For any i, let $\tilde{v}_i$ be a corresponding flux-path for v_i. Then:*

$$v = \sum_{1\leq i\leq k} a_i \sum \tilde{v}_i.$$

Proof. We directly have $v = \sum_{1\leq i\leq k} a_i v_i$ and for any i, $v_i = \sum \tilde{v}_i$. □

We can now prove the key lemma of this section.

Lemma A2. *Let W, W_0 be reaction networks such that $W_0 \in LinNets$ and $W_0 \Rrightarrow_c^* W$. Then the following properties hold:*

1. *for any flux $v;e \in W$, there is a corresponding flux-path $\tilde{v}$ for W such that $\sum \tilde{v} = v$. Moreover, if v is not dependent then, for any intermediate species X, we have $\tilde{v}(X) \leq 1$;*
2. *for any flux-path $\tilde{v}$ for W such that for any X, $\tilde{v}(X) \leq 1$, there is a corresponding flux $v;e \in W$, that is $\sum \tilde{v} = v$;*
3. *if $v;e \in W$ depends on $v_1;e_1,\dots,v_k;e_k \in W$, then*
   - *there exists an index i such that $Cons_{\mathcal{I}}(\boldsymbol{r}_{v_i}) = Cons_{\mathcal{I}}(\boldsymbol{r}_v)$, $Prod_{\mathcal{I}}(\boldsymbol{r}_{v_i}) = Prod_{\mathcal{I}}(\boldsymbol{r}_v)$, and*
   - *for any $j \neq i$, $Prod_{\mathcal{I}}(\boldsymbol{r}_{v_j}) = Cons_{\mathcal{I}}(\boldsymbol{r}_{v_j}) = \emptyset$ and any flux-path $\tilde{v}_j$ that corresponds to v is circular.*

Proof. We proceed by induction on the simplification steps. We start by proving each conclusion of the Lemma for the base case, that is $W = W_0$.

(1) for flux $v;e$ in the initial network W_0, v is necessarily a unary vector v_i for some i. So we can directly associate the flux-path $\tilde{v} = v_i$ that trivially corresponds to v. Since it is a unary vector, the flux v is also necessarily not dependent. Because a flux-path of size 1 is always non-circular, we also have, for any intermediate species X, $\tilde{v}(X) = 0$, and thus $\tilde{v}(X) \leq 1$ as required.

(2) any flux-path $\tilde{v}$ for W_0 is necessarily of size 1. Otherwise, for $\tilde{v}$ being a non-circular flux-path, there would exist some $X \in Specs(W_0)\backslash Specs(W_0) = \emptyset$. And for $\tilde{v}$ being a circular flux-path, there would exist at least two species X and Y such that $\tilde{v}(X) > 0$ and $\tilde{v}(Y) > 0$, which contradicts the definition of circular flux-path. Then there exists $v_i;e \in W_0$ such that $v_i = \tilde{v}$.

(3) as said above, a flux $v;e \in W_0$ v can not be dependent.

Now, considering the inductive case, we assume that the Lemma is true for a network W' such that $W_0 \Rrightarrow_c^k W'$ (for some $k > 0$) and $W' \Rrightarrow_c W$. If $W' \Rrightarrow_{\text{C-MOD}} W$ or $W' \Rrightarrow_{\text{C-SIM}} W$, only the kinetics are modified between W' and W. Therefore, the Lemma is still true in W. It remains to investigate the cases $W' \Rrightarrow_{\text{C-DEP}} W$ and $W' \Rrightarrow_{\text{C-INTER}} W$.

(C-DEP) Assuming that $W' \Rrightarrow_{\text{C-DEP}} W$, we prove that each point of the Lemma is satisfied by W.

(1) Let $v;e \in W$, then it is the case that $v;e' \in W'$ for some expression e' because the rule (C-DEP) only removes a dependent flux and modifies some kinetic expressions. By induction hypothesis, there is a flux-path $\tilde{v}$ for W' such that $\sum \tilde{v} = v$. Also, because $Specs(W) = Specs(W')$, any flux-path for W' is also a flux-path for W, which proves that $\tilde{v}$ is a flux-path for v. Finally, if v is dependent in W, it is necessarily dependent in W and satisfies $\forall X \in \mathcal{I}.\tilde{v}(X) \leq 1$ by induction hypothesis.

(2) Let $\tilde{v}$ be a flux-path for W such that, for any intermediate species X, $\tilde{v}(X) \leq 1$. Again, because $Specs(W) = Specs(W')$, $\tilde{v}$ is also a flux-path for W'. By induction hypothesis, there is a corresponding flux $v;e \in W'$. If $v;e$ is not the flux that is removed by the application of (C-DEP), then this flux still occurs in W (possibly with an updated kinetic) and we conclude directly. We now show that it can not actually be otherwise, and more precisely, that assuming v removed by (C-DEP) contradicts $\tilde{v}(X) \leq 1$.

Suppose that v is removed by (C-DEP), then v depends on some fluxes $v_1; e_1, \dots v_k; e_k$ in W'. For the sake of simplicity, we only consider the case where $Cons_{\mathcal{I}}(\mathbf{r}_{\tilde{v}}) = X$ for some intermediate $X \in W$. The other case works similarly. We necessarily have $k > 1$, because $k = 1$ contradicts the linearity assumption. By induction hypothesis and Point 3, there exists in particular at least one j such that $Prod_{\mathcal{I}}(\mathbf{r}_{v_j}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_j}) = \emptyset$. Again, by induction hypothesis, there exists $\tilde{v'} = v'_1, \dots, v'_l$ a circular flux-path corresponding to v_j in W'. Since it is circular, there exist some intermediate species $X_1, \dots, X_l$ such that for any $1 \leq i < l$, $Cons_{\mathcal{I}}(\mathbf{r}_{v'_i}) = X_i$ and $Prod_{\mathcal{I}}(\mathbf{r}_{v'_i}) = X_{i+1}$, and $Cons_{\mathcal{I}}(\mathbf{r}_{v'_l}) = X_l$, $Prod_{\mathcal{I}}(\mathbf{r}_{v'_l}) = X_1$. Note that we also have $X_i \neq X$, since, by definition of flux-path, $X_i \notin W'$ for any i, while $X \in W$. Since $\sum \tilde{v} = v = \sum v_k$ and $v_j = \sum \tilde{v'} = \sum_i v'_i$, the unit vectors v'_i also appear in the flux-path $\tilde{v}$. So the intermediate species X_i are present at least one time in $\tilde{v}$. Moreover, there is at least one i such that in $\tilde{v}$, the unit vector v_i is preceded by a unit vector that is not one of $\tilde{v'}$. Therefore, X_i is produced by another flux, i.e., $\tilde{v}(X) > 1$, which contradicts the hypothesis.

(3) Let $v; e \in W$ be a flux dependent on $v_1; e_1, \dots, v_n; e_n \in W$, that is, in particular, $v = \sum_{1 \leq i \leq k} n_i v_i$ for some $n_i > 0$. Since (C-DEP) removes one flux and possibly modifies some kinetics, there is a flux $v; e' \in W$ in W' that either depends on $v_1; e'_1, \dots, v_n; e'_n \in W$ or on $v_0; e'_0, v_1; e'_1, \dots, v_n; e'_n \in W$ where $v_0; e'_0$ is the flux removed by (C-DEP). The latter case is not possible, since it would imply that $v = \sum_{1 \leq i \leq k} n_i v_i = n_0 v_0 + \sum_{1 \leq i \leq k} n_i v_i$ for some $n_0 > 0$ and unary vector v_0. Thus, we conclude that $v; e' \in W$ depends on $v_1; e'_1, \dots, v_n; e'_n \in W$ that, by induction hypothesis, satisfies the conditions of Point 3.

(C-INTER) Now, assuming that $W' \Rrightarrow_{\text{C-INTER}} W$, we again prove that each point of the Lemma is satisfied by W.

(1) Let $v; e$ be in W. Either there is a corresponding flux $v; e' \in W'$, and we conclude directly by induction hypothesis, or $v; e$ is the result of merging some $v_p; e_p \in W'$ that produces X and some $v_c; e_c \in W'$ that consume it. In this case, by induction hypothesis, there are some corresponding flux-paths $\tilde{v_p}$ and $\tilde{v_c}$. The concatenation $\tilde{v} = \tilde{v_p}\tilde{v_c}$ of these paths is a flux-path. Indeed,

- the production of $\tilde{v_p}$ coincides with the consumption of $\tilde{v_c}$ because there is an intermediate species X such that $Prod_{\mathcal{I}}(\mathbf{r}_{v_p}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_c}) = \{X\}$, $Prod_{\mathcal{I}}(\mathbf{r}_{\tilde{v_p}}) = Prod_{\mathcal{I}}(\mathbf{r}_{v_p})$ and $Cons_{\mathcal{I}}(\mathbf{r}_{\tilde{v_c}}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_c})$.
- if $\tilde{v}$ is non-circular, for any intermediate species Y such that $\tilde{v}(Y) > 0$, either $Y = X$ and $Y \in Specs(W_0) \backslash Specs(W)$, or, $\tilde{v_p}(Y) > 0$ or $\tilde{v_c}(Y) > 0$ and by induction hypothesis, $Y \in Specs(W_0) \backslash Specs(W')$ that is $Y \in Specs(W_0) \backslash Specs(W)$.
- if $\tilde{v}$ is circular, there exists an intermediate species Y which is both consumed by v_p and produced by v_c. The flux-path $\tilde{v_c}$ cannot be circular, as this would imply $Cons_{\mathcal{I}}(\mathbf{r}_{v_c}) = Prod_{\mathcal{I}}(\mathbf{r}_{v_c}) = \emptyset \neq X$, and similarly for $\tilde{v_p}$. By definition of non-circular flux-path, X and Y are the only two species in $\tilde{v_c}$ and $\tilde{v_p}$ such that $X \in Specs(W')$ and $Y \in Specs(W')$. Thus, Y is the only non-eliminated intermediate species in $\tilde{v}$ w.r.t. W, meaning again that $\tilde{v}$ is indeed a circular flux-path.

Moreover, $\tilde{v}$ trivially corresponds to v. We prove with Point 3 that if $\tilde{v}$ is non-dependent, then $\tilde{v}(X) \leq 1$ for any X.

(2) Let $\tilde{v}$ be a flux-path for W such that for any Y, $\tilde{v}(Y) \leq 1$. Let X be the intermediate species removed by (C-INTER), in particular $\tilde{v}(X) \leq 1$, hence either $\tilde{v}(X) = 0$ or $\tilde{v}(X) = 1$. Since $Specs(W) = Specs(W') \backslash \{X\}$, if $\tilde{v}(X) = 0$, then $\tilde{v}$ is also a flux-path for W', so by induction there is a corresponding flux $v; e' \in W'$. Since $\tilde{v}(X) = 0$, we still have a flux $v; e \in W$, that corresponds to $\tilde{v}$. If $\tilde{v}(X) = 1$ then we can decompose $\tilde{v}$ into $\tilde{v_p}$ producing X and $\tilde{v_c}$ consuming X such that $\tilde{v} = \tilde{v_p}\tilde{v_c}$. $\tilde{v_p}$ and $\tilde{v_c}$ can not be circular, as this would imply that X is both consumed and produced by $\tilde{v_p}$ and by $\tilde{v_c}$, contradicting the fact that $\tilde{v}(X) = 1$. Therefore $\tilde{v_p}(X) = \tilde{v_c}(X) = 0$. Again, because $Specs(W) = Specs(W') \backslash \{X\}$ and X is the species removed

by (C-INTER) and $\tilde{v}$ is a flux-path, $\tilde{v_p}$ and $\tilde{v_c}$ are (non-circular) flux-paths for W'. We can then apply the induction hypothesis and infer that there are some corresponding fluxes $v_p, v_c \in W'$, the first one that produces X and the second one that consumes it. Consequently, there is a flux $v;e \in W$ that is the merging of v_p and v_c, and that corresponds to $\tilde{v}$.

(3) Let $v;e \in W$ be a flux that depends on $v_1;e_1,\ldots,v_k;e_k \in W$ and X be the intermediate species removed by (C-INTER). We distinguish two cases: either (case 1) $v;e$ is the simplification of some flux $v;e'$ (meaning that v does neither produce nor consume X) or (case 2) it results from merging fluxes that produce and consume X.

(Case 1) By induction hypothesis and Point 1, there exists a flux-path $\tilde{v}$ corresponding to v for W'. We have $\tilde{v}(X) = 0$ since X has not been removed. Suppose that there is $i \in \{1,\ldots,k\}$ such that v_i is the merging, by (C-INTER), of fluxes that produce and a consume X. In this case, for any $\tilde{v_i}$ corresponding to v_i, $\tilde{v_i}(X) > 0$ and, by the Lemma A1, we would have that $\tilde{v}(X) > 0$, which contradicts $\tilde{v}(X) = 0$. Therefore, none of the v_is are the merging of other fluxes by (C-INTER), therefore, there are $v_1;e'_1,\ldots,v_k;e'_k \in W'$ such that $v;e'$ depends on those fluxes. Since, by induction hypothesis, Point 3 is satisfied for $v;e'$ in W', it is also satisfied for $v;e$ in W.
(Case 2) Let $v_p;e_p \in W'$ be a flux that produces X, and $v_c;e_c \in W'$ a flux that consumes it and $v;e \in W$ their merging. Let $\tilde{v_p}$ and $\tilde{v_c}$ be flux-paths in W' for, respectively, v_p and v_c (such flux paths exist by induction hypothesis). Any corresponding path $\tilde{v}$ is then the concatenation of corresponding paths $\tilde{v_p}$ and $\tilde{v_c}$.

We first prove that, if either $v_p;e_p$ or $v_c;e_c$ is dependent in W', then $v;e$ is also dependent in W. By induction, if $v_p;e_p$ depends on $v'_1;e'_1,\ldots,v'_\ell;e'_\ell$, there exists a unique i such that $Cons_{\mathcal{I}}(\mathbf{r}_{v'_i}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_p})$, $Prod_{\mathcal{I}}(\mathbf{r}_{v'_i}) = Prod_{\mathcal{I}}(\mathbf{r}_{v_p}) = \{X\}$, and, for any $j \neq i$, $Cons_{\mathcal{I}}(\mathbf{r}_{v'_j}) = Cons_{\mathcal{I}}(\mathbf{r}_{v'_j}) = \emptyset$. Then, there is a flux $v''_i;e''_i$ in W that is the merging of v'_i and v_c, and there are fluxes $v'_j;e'_j$ in W for $j \neq i$. Then $v;e$ depends on $v''_i;e''_i$ and the $v'_j;e'_j$.

If both $v_p;e_p$ and $v_c;e_c$ are not dependent, by induction hypothesis, for any $Y \neq X$, in the corresponding flux-path, we have $\tilde{v_p}(Y) \leq 1$ and $\tilde{v_c}(Y) \leq 1$. If $\tilde{v_p}(Y) + \tilde{v_c}(Y) \leq 1$, then $\tilde{v}(Y) \leq 1$. We also have $\tilde{v}(X) = 1$ (indeed, X is the removed species, so $\tilde{v_p}(X) = \tilde{v_c}(X) = 0$ and the X produced by $\tilde{v_p}$ is merged with the X consumed by $\tilde{v_c}$ in $\tilde{v}$). Therefore, in this case Point 1 is satisfied. If there is a Y such that $\tilde{v_p}(Y) = \tilde{v_c}(Y) = 1$, then Y occurs twice in $\tilde{v}$, and there is an intermediate species Z (that can possibly be Y if no other intermediate species occurs more than once between both occurrences of Y) such that $\tilde{v}(Z) = 2$ and such that there is a circular flux-path $\widetilde{v_{cyc}}$, subpath of $\tilde{v}$ that begins and end with Z:

$$\tilde{v} : \overbrace{\underbrace{\ldots \to Y \to \ldots}_{\tilde{v_1}} \to \underbrace{Z \to \ldots \to X}_{\widetilde{v_{cyc}}}}^{\widetilde{v_p}} \to \overbrace{\underbrace{\ldots \to Z}_{\widetilde{v_{cyc}}} \to \underbrace{\ldots \to Y \to \ldots}_{\tilde{v_2}}}^{\tilde{v_c}}$$

Then using Point 2, there is a corresponding flux $v_{cyc} \in W$, with $Cons_{\mathcal{I}}(\mathbf{r}_{v_{cyc}}) = Prod_{\mathcal{I}}(\mathbf{r}_{v_{cyc}})$. We can repeat the same operation on the remaining path $\tilde{v_1}\tilde{v_2}$, and obtain at each step a new (circular) flux. We stop when we obtain a remaining path $v_{\tilde{rem}}$, such that for any Y, we have $v_{\tilde{rem}}(Y) \leq 1$. Then there is a corresponding flux v_{rem} with $Cons_{\mathcal{I}}(\mathbf{r}_{v_{rem}}) = Cons_{\mathcal{I}}(\mathbf{r}_v)$, and $Prod_{\mathcal{I}}(\mathbf{r}_{v_{rem}}) = Prod_{\mathcal{I}}(\mathbf{r}_v)$. Then v is dependent on v_{rem} and the set of circular fluxes we obtained in this process. Therefore, in this case Point 3 is satisfied.

Now assume that $v_p;e_p$ and $v_c;e_c$ are dependent, and so that $v;e$ is dependent too. By induction and using Point 3, $v_p;e_p$ depends on a flux $v_{p,i};e_{p,i}$ with $Cons_{\mathcal{I}}(\mathbf{r}_{v_{p,i}}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_p})$ and $Prod_{\mathcal{I}}(\mathbf{r}_{v_{p,i}}) = Prod_{\mathcal{I}}(\mathbf{r}_{v_p})$, and on other fluxes such that $Cons_{\mathcal{I}}(\mathbf{r}_{v_{p,j}}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_p}) = \emptyset$, and similarly for $v_c;e_c$. Then any $v_{p,j};e_{p,j}$ and any $v_{c,j};e_{c,j}$ is still in W, while $v_{p,i};e_{p,i}$ is merged with $v_{c,i};e_{c,i}$, forming a new flux $v_i;e_i$. Then $v;e$ depends on $v_i;e_i$, the $v_{p,j};e_{p,j}$ and the $v_{c,j};e_{c,j}$. Since $Prod_{\mathcal{I}}(\mathbf{r}_{v_i}) = Prod_{\mathcal{I}}(\mathbf{r}_{v_{c,i}}) = Prod_{\mathcal{I}}(\mathbf{r}_{v_c}) = Prod_{\mathcal{I}}(\mathbf{r}_v)$, and similarly $Cons_{\mathcal{I}}(\mathbf{r}_{v_i}) = Cons_{\mathcal{I}}(\mathbf{r}_v)$, Point 3 is satisfied.

□

Then Proposition 3 is a direct corollary of Point 3 of Lemma A2.

Appendix C. Proofs of the Full Confluence of the Simplification

We prove Lemmas 9 and 12.

Lemma A3. *Let W be a network such that $W \Rrightarrow_{\text{C-DEP}} W_1$ and $W \Rrightarrow_{\text{C-INTER}} W_2$. Then $\exists W_i'$ such that $W_i \Rrightarrow_C^* W_i'$ and $W_1' \cong W_2'$.*

Proof. Let X be the intermediate species, and $v_d; e_d$ the dependent reaction, that depends on $v_1; e_1, \dots, v_n; e_n$, with coefficients $a_1, \dots, a_n$.

The main idea is to use the Proposition 3 to prove that if X is in the dependent flux v_d, it is also in one of the fluxes v_i whose v_d depends on. Therefore, if we eliminate X and combine v_d with another flux v', we also merge v_i with v'. Then $v_d \diamond_X v'$ is still dependent on $v_i \diamond_X v'$ and other fluxes, and can be removed.

Let us first assume that X is not involved in v_d (that is, $X \notin Prod_{\mathcal{I}}(\mathbf{r}_{v_d}) \cup Cons_{\mathcal{I}}(\mathbf{r}_{v_d})$). By Lemma A2, Point 3, X is not involved in the v_i either. Then v_d will still be dependent in W_2, so we can still remove it after removing X. Reciprocally, X can still be removed in W_1. We now show that the simplification of the fluxes v_i, e_i are the same in W_1' and W_2'. The case for the other fluxes is trivial. In W_1', we obtain the flux $v_i; (e_i + a_i e_d)[X := X(0)]$. In W_2', we obtain $v_i; e_i[X := X(0)] + a_i e_d[X := X(0)]$. Therefore, these two expressions are similar, and $W_1' \cong W_2'$.

Assume that $X \in v_d$, for instance $Prod_{\mathcal{I}}(\mathbf{r}_{v_d}) = \{X\}$. Then, again by Lemma A2 and Point 3, there is a flux $v_i; e_i$ with $Prod_{\mathcal{I}}(\mathbf{r}_{v_i}) = \{X\}$, and for any other $j \neq i$, $Prod_{\mathcal{I}}(\mathbf{r}_{v_j}) = Cons_{\mathcal{I}}(\mathbf{r}_{v_j}) = \emptyset$. We denote by *prod* and *cons* the expressions as defined in the rule (C-INTER), by V_{prod} the fluxes that produce some X, and V_{cons} the ones that consume it.

In W_2 after removing X, we obtain the fluxes:

- the combination of v_d and the consuming fluxes: $\{v_d \diamond v_{cons}; e_d e_{cons} / cons \mid v_{cons}; e_{cons} \in V_{cons}\}$,
- the combination of r_i and the consuming reactions: $\{v_i \diamond v_{cons}; e_i e_{cons} / cons \mid v_{cons}; e_{cons} \in V_{cons}\}$,
- the other combined fluxes: $\{v_{prod} \diamond v_{cons}; e_{prod} e_{cons} / cons \mid v_{prod}; e_{prod} \in V_{prod}, v_{cons}; e_{cons} \in V_{cons}\}$,
- the remaining fluxes not combined: $\{v_j; e_j[X := X(0)]\}_{j \neq i}$,
- the other fluxes that are not in V_{prod}, V_{cons}, v_j, where we substitute X by $X(0)$.

Since v_d was dependent on $v_1, \dots, v_n$, we have that any flux $v_d \diamond v_{cons}$ in the first set is dependent on a flux $v_i \diamond v_{cons}$ in the second set and the fluxes v_j. Therefore, we can recursively remove those fluxes with the rule (C-DEP). We obtain the network W_2' with the fluxes:

- $\{v_i \diamond v_{cons}; e_i e_{cons} / cons + e_d e_{cons} / cons\}$,
- $\{v_{prod} \diamond v_{cons}; e_{prod} e_{cons} / cons\}$,
- $\{v_j; e_j)[X := X(0)] + \sum_{e_{cons}} a_i e_d e_{cons} / cons\}$,
- the other fluxes that are not in V_{prod}, V_{cons}, v_j, where we substitute X by $X(0)$.

Now, if we first remove v_d, in W_1 we obtain the fluxes:

- $v_i; e_i + e_d$
- $\{v_j; e_j + a_j e_d\}_{j \neq i}$
- $V_{cons} \setminus \{v_d, v_i\}$,
- V_{prod},
- the other fluxes that are not in V_{prod}, V_{cons}, v_j.

We can still remove X, we obtain the reactions:

- $\{v_i \diamond v_{cons}; (e_i + e_d) e_{cons} / cons\}$,

- $\{v_{prod} \diamond v_{cons}; e_{prod}e_{cons}/cons\}$,
- $\{v_j; (e_j + a_j e_d)[X := X(0)]\}$,
- the other fluxes that are not in V_{prod}, V_{cons}, v_j, where we substitute X by $X(0)$.

Note that we have:

$$\begin{aligned} e_j[X := X(0)] + \sum_{e_{cons}} a_j e_d e_{cons}/cons &\cong e_j[X := X(0)] + a_j e_d \\ &\cong (e_j + a_j e_d)[X := X(0)]. \end{aligned}$$

The fluxes of the two simplified networks are therefore similar, that is $W_1' \cong W_2'$. The case with $Cons_{\mathcal{I}}(\mathbf{r}_{v_d}) = \{X\}$ is similar. □

Lemma A4. *Let W be a network such that $W \Rightarrow_{\text{C-INTER}} W_i$ for $i \in \{1,2\}$. Then $\exists W_i'$ such that $W_i \Rightarrow_c^* W_i'$ and $W_1' \cong W_2'$.*

Proof. The main idea here is that after removing one intermediate species, we can still remove the other one, either with (C-MOD) or with (C-INTER). In the second case, some dependent fluxes are generated, that we can eliminate to find the same simplified network, whatever the order of elimination of the intermediate species.

Let X and Y be the intermediate species removed to obtain W_1 and W_2. We can partition the fluxes of W into:

- $V_X = \{v_X; e_X \mid X \in Prod_{\mathcal{I}}(\mathbf{r}_{v_X}), Y \notin v_X\}$, the fluxes producing X without Y,
- $V_{X'} = \{v_{X'}; Xe_{X'} \mid X \in Cons_{\mathcal{I}}(\mathbf{r}_{v_{X'}}), Y \notin v_{X'}\}$, the fluxes consuming X without Y,
- $V_{mod(X)} = \{v_{mod(X)}; e_{mod(X)} \mid X \notin Prod_{\mathcal{I}}(\mathbf{r}_{v_{mod(X)}}) \cup Cons_{\mathcal{I}}(\mathbf{r}_{v_{mod(X)}}), X \in Vars(e_{mod(X)}), Y \notin v_{mod(X)}\}$, the fluxes with modifier X and without Y,
- $V_Y = \{v_Y; e_Y \mid Y \in Prod_{\mathcal{I}}(\mathbf{r}_{v_Y}), X \notin v_Y\}$, the fluxes producing Y without X,
- $V_{Y'} = \{v_Y; Ye_{Y'} \mid Y \in Cons_{\mathcal{I}}(\mathbf{r}_{v_Y}), X \notin v_Y\}$, the fluxes consuming Y without X,
- $V_{mod(Y)} = \{v_{mod(Y)}; e_{mod(Y)} \mid Y \notin Prod_{\mathcal{I}}(\mathbf{r}_{v_{mod(Y)}}) \cup Cons_{\mathcal{I}}(\mathbf{r}_{v_{mod(Y)}}), Y \in Vars(e_{mod(Y)}), X \notin v_{mod(Y)}\}$, the fluxes with modifier Y and without X,
- $V_{XY'} = \{v_{XY'}; Ye_{XY'} \mid X \in Prod_{\mathcal{I}}(\mathbf{r}_{v_{XY'}}), Y \in Cons_{\mathcal{I}}(\mathbf{r}_{r_{XY'}})\}$, the fluxes producing X and consuming Y,
- $V_{X'Y} = \{v_{X'Y}; Xe_{X'Y} \mid Y \in Prod_{\mathcal{I}}(\mathbf{r}_{v_{X'Y}}), X \in Cons_{\mathcal{I}}(\mathbf{r}_{v_{X'Y}})\}$, the fluxes producing Y and consuming X,
- $V_{mod(XY)} = \{v_{mod(XY)}; e_{mod(XY)} \mid X, Y \notin v_{mod(XY)}\}$, the fluxes with modifier X and Y.

We define the following variables:

- $T_X = \sum_{V_X} e_X$ $\qquad$ $T_{X'} = \sum_{V_{X'}} e_{X'}$
- $T_Y = \sum_{V_Y} e_Y$ $\qquad$ $T_{Y'} = \sum_{V_{Y'}} e_{Y'}$
- $T_{X'Y} = \sum_{V_{X'Y}} e_{X'Y}$ $\qquad$ $T_{XY'} = \sum_{V_{XY'}} e_{XY'}$

Let first remove X. We obtain the following combined fluxes:

- $V_X \diamond V_{X'} = \{v_X \diamond v_{X'}; e_X e_{X'}/(T_{X'} + T_{X'Y})\}$,
- $V_X \diamond V_{X'Y} = \{v_X \diamond v_{X'Y}; e_X e_{X'Y}/(T_{X'} + T_{X'Y})\}$,
- $V_{XY'} \diamond V_{X'} = \{v_{XY'} \diamond v_{X'}; e_{XY'} e_{X'} Y/(T_{X'} + T_{X'Y})\}$,
- $V_{XY'} \diamond V_{X'Y} = \{v_{XY'} \diamond v_{X'Y}; e_{XY'} e_{X'Y} Y/(T_{X'} + T_{X'Y})\}$.

The fluxes with X as modifier become (with $X(0) = (T_X + YT_{XY'})/(T_{X'} + T_{X'Y})$:

- $V'_{mod(X)} = \{v_{mod(X)}; e_{mod(X)}[X := X(0)]\}$,
- $V'_{mod(XY)} = \{v_{mod(XY)}; e_{mod(XY)}[X := X(0)]\}$.

Finally, some fluxes are not modified:

- $V_Y = \{v_Y; e_Y\}$,
- $V_{Y'} = \{v_Y; Ye_{Y'}\}$,
- $V_{mod(Y)} = \{v_{mod(Y)}; e_{mod(Y)}\}$.

There are now two cases to consider. First, it is possible that Y is now only a modifier in W_1. This means that $V_X = V_{X'} = V_Y = V_{Y'} = 0$, that is any flux with X as reactant (resp. product) also admits Y as product (resp. reactant), and reciprocally. We can then apply (C-MOD) on W_1, and obtain the fluxes (with $X(0) = Y(0)T_{XY'}/T_{X'Y}$):

- $V_{XY'} \diamond V_{X'Y} = \{v_{XY'} \diamond v_{X'Y}; e_{XY'}e_{X'Y}Y(0)/T_{X'Y}\}$.
- $V'_{mod(X)} = \{v_{mod(X)}; e_{mod(X)}[X := X(0)]\}$,
- $V_{mod(Y)} = \{v_{mod(Y)}; e_{mod(Y)}[Y := Y(0)]\}$,
- $V'_{mod(XY)} = \{v_{mod(XY)}; e_{mod(XY)}[x_X := X(0)][Y := Y(0)]\}$.

Using the constraint $X(0) = Y(0)T_{XY'}/T_{X'Y}$ to rewrite the first kinetic expression into $e_{XY'}e_{X'Y}X(0)/T_{XY'}$, we can see by symmetry that we obtain similar fluxes by removing Y first (with (C-INTER)) and then X (with (C-MOD)).

In the other case, we can still remove Y with (C-INTER). We first compute the sum U_Y (resp. $U_{Y'}$) of the kinetics of the fluxes that produced (resp. consumed) Y :

$$U_Y = \frac{T_{X'}T_Y + T_YT_{X'Y} + T_XT_{X'Y}}{T_{X'} + T_{X'Y}} \qquad U_{Y'} = \frac{T_{X'}T_{Y'} + T_{Y'}T_{X'Y} + T_{X'}T_{XY'}}{T_{X'} + T_{X'Y}}.$$

We write $T = T_{X'}T_{Y'} + T_{Y'}T_{X'Y} + T_{X'}T_{XY'}$. We obtain the following combined fluxes:

- $(V_X \diamond V_{X'Y}) \diamond (V_{XY'} \diamond V_{X'}) = \{v_X \diamond v_{X'Y} \diamond v_{X'} \diamond v_{XY'}; \frac{e_Xe_{X'}e_{XY'}e_{X'Y}}{(T_{X'} + T_{X'Y})T}\}$,
- $(V_X \diamond V_{X'Y}) \diamond V_{Y'} = \{v_X \diamond v_{X'Y} \diamond v_{Y'}; \frac{e_Xe_{Y'}e_{X'Y}}{T}\}$,
- $V_Y \diamond (V_{XY'} \diamond V_{X'}) = \{v_Y \diamond v_{XY'} \diamond v_{X'}; \frac{e_Ye_{XY'}e_{X'}}{T}\}$,
- $V_Y \diamond V_{Y'} = \{v_Y \diamond v_{Y'}; \frac{e_Ye_{Y'}(T_{X'} + T_{X'Y})}{T}\}$.

The fluxes with Y as modifier become:

- $(V_{XY'} \diamond V_{X'Y})' = \{v_{XY'} \diamond v_{X'Y}; \frac{T_{X'}T_Y + T_YT_{X'Y} + T_XT_{X'Y}}{T(T_{X'} + T_{X'Y})}e_{XY'}e_{X'Y}\}$,
- $V''_{mod(X)} = \{v_{mod(X)}; e_{mod(X)}[X := \frac{T_XT_{Y'} + T_XT_{XY'} + T_YT_{XY'}}{T}]\}$,
- $V''_{mod(XY)} = \{v_{mod(XY)}; e_{mod(XY)}[X := \frac{T_XT_{Y'} + T_XT_{XY'} + T_YT_{XY'}}{T}][Y := \frac{T_YT_{X'} + T_YT_{X'Y} + T_XT_{X'Y}}{T}]\}$,
- $V'_{mod(Y)} = \{v_{mod(Y)}; e_{mod(Y)}[x_Y := \frac{T_YT_{X'} + T_YT_{X'Y} + T_XT_{X'Y}}{T}]\}$.

Finally, some fluxes do not involve Y:

- $R_X \diamond R_{X'} = \{vec(r_X) + vec(r_{X'}); e_Xe_{X'}/(T_{X'} + T_{X'Y})\}$

Now we can observe that we obtained some dependent fluxes. Any flux in $(V_X \diamond V_{X'Y}) \diamond (V_{XY'} \diamond V_{X'})$ is the combination of a flux from $(V_{XY'} \diamond V_{X'Y})'$ and $V_X \diamond V_{X'}$. We can then remove the first flux, while modifying the kinetics of the others. Finally, after simplifying the kinetic expressions by similarity, the simplified network has the following fluxes:

- $V_X \diamond V_{X'} = \{v_X \diamond v_{X'}; \frac{e_Xe_{X'}(T_{Y'} + T_{XY'})}{T}\}$,
- $V_Y \diamond V_{Y'} = \{v_Y \diamond v_{Y'}; \frac{e_Ye_{Y'}(T_{X'} + T_{X'Y})}{T}\}$,

- $(V_X \diamond V_{X'Y}) \diamond V_{Y'} = \{v_X \diamond v_{X'Y} \diamond v_{Y'}; \frac{e_X e_{Y'} e_{X'Y}}{T}\}$,
- $V_Y \diamond (V_{XY'} \diamond V_{X'}) = \{v_Y \diamond v_{XY'} \diamond v_{X'}; \frac{e_Y e_{XY'} e_{X'}}{T}\}$,
- $V''_{mod(X)} = \{v_{mod(X)}; e_{mod(X)}[x_X := \frac{T_X T_{Y'} + T_X T_{XY'} + T_Y T_{XY'}}{T}]\}$,
- $V'_{mod(Y)} = \{v_{mod(Y)}; e_{mod(Y)}[x_Y := \frac{T_Y T_{X'} + T_Y T_{X'Y} + T_X T_{X'Y}}{T}]\}$,
- $(V_{XY'} \diamond V_{X'Y})' = \{v_{XY'} \diamond v_{X'Y}; \frac{(T_Y + T_X) e_{XY'} e_{X'Y}}{T}\}$,
- $V''_{mod(XY)} = \{v_{mod(XY)}; e_{mod(XY)}[x_X := \frac{T_X T_{Y'} + T_X T_{XY'} + T_Y T_{XY'}}{T}][x_Y := \frac{T_Y T_{X'} + T_Y T_{X'Y} + T_X T_{X'Y}}{T}]\}$,

We can observe that:

- the 2 first sets are symmetric to each other, in the sense that if we switch X and Y in the first set, we obtain the second one,
- the 2 following sets are symmetric to each other,
- the 2 following sets are symmetric to each other too,
- the following set is symmetric in X and Y,
- the last set is symmetric in X and Y (since the substitutions commute).

Therefore, by symmetry, we can obtain exactly the same network if we first remove Y, then remove X and the dependent fluxes. We conclude that $W'_1 \cong W'_2$.

□

References

1. Feinberg, M. Chemical reaction network structure and the stability of complex isothermal reactors—I. The deficiency zero and deficiency one theorems. *Chem. Eng. Sci.* **1987**, *42*, 2229–2268.
2. Hucka, M.; Finney, A.; Sauro, H.M.; Bolouri, H.; Doyle, J.C.; Kitano, H.; Arkin, A.P.; Bornstein, B.J.; Bray, D.; Cornish-Bowden, A.; et al. The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics* **2003**, *19*, 524–531.
3. Calzone, L.; Fages, F.; Soliman, S. BIOCHAM: An environment for modeling biological systems and formalizing experimental knowledge. *Bioinformatics* **2006**, *22*, 1805–1807.
4. Kuttler, C.; Lhoussaine, C.; Nebut, M. Rule-based modeling of transcriptional attenuation at the tryptophan operon. In *Transactions on Computational Systems Biology XII*; Springer: New York, NY, USA, 2010; pp. 199–228.
5. Chaouiya, C. Petri net modelling of biological networks. *Brief. Bioinform.* **2007**, *8*, 210–219.
6. Juty, N.; Ali, R.; Glont, M.; Keating, S.; Rodriguez, N.; Swat, M.J.; Wimalaratne, S.M.; Hermjakob, H.; Le Novère, N.; Laibe, C.; et al. BioModels: Content, Features, Functionality and Use. *CPT Pharmacomet. Syst. Pharmacol.* **2015**, *4*, 55–68.
7. Mäder, U.; Schmeisky, A.G.; Flórez, L.A.; Stülke, J. SubtiWiki—A comprehensive community resource for the model organism *Bacillus subtilis*. *Nucleic Acids Res.* **2012**, *40*, 1278–1287.
8. Niehren, J.; John, M.; Versari, C.; Coutte, F.; Jacques, P. Qualitative Reasoning about Reaction Networks with Partial Kinetic Information. In *Computational Methods for Systems Biology*; Lecture Notes in Computer Science; Springer: New York, NY, USA, 2015; Volume 9308, pp. 157–169.
9. Radulescu, O.; Gorban, A.N.; Zinovyev, A.; Noel, V. Reduction of dynamical biochemical reactions networks in computational biology. *Front. Genet.* **2012**. *3*, 131.
10. Michaelis, L.; Menten, M.L. Die kinetik der invertinwirkung. *Biochem. Z.* **1913**, *49*, 333–369. (In German)
11. Segel, L.A. On the validity of the steady state assumption of enzyme kinetics. *Bull. Math. Biol.* **1988**, *50*, 579–593.
12. Cornish-Bowden, A. *Fundamentals of Enzyme Kinetics*; Wiley: Hoboken, NJ, USA, 2013.
13. Heineken, F.; Tsuchiya, H.; Aris, R. On the mathematical status of the pseudo-steady state hypothesis of biochemical kinetics. *Math. Biosci.* **1967**, *1*, 95–113.

14. Segel, L.A.; Slemrod, M. The quasi-steady-state assumption: A case study in perturbation. *SIAM Rev.* **1989**, *31*, 446–477.
15. King, E.L.; Altman, C. A schematic method of deriving the rate laws for enzyme-catalyzed reactions. *J. Phys. Chem.* **1956**, *60*, 1375–1378.
16. Chou, K.-C.; Forsen, S. Graphical rules of steady-state reaction systems. *Can. J. Chem.* **1981**, *59*, 737–755.
17. Fages, F.; Gay, S.; Soliman, S. Inferring reaction systems from ordinary differential equations. *Theor. Comput. Sci.* **2015**, *599*, 64–78.
18. Sáez, M.; Wiuf, C.; Feliu, E. Graphical reduction of reaction networks by linear elimination of species. *J. Math. Biol.* **2017**, *74*, 195–237.
19. Madelaine, G.; Lhoussaine, C.; Niehren, J. Attractor Equivalence: An Observational Semantics for Reaction Networks. In *Formal Methods in Macro-Biology*; Springer: New York, NY, USA, 2014.
20. Schmidt-Schauss, M.; Sabel, D.; Niehren, J.; Schwinghammer, J. Observational Program Calculi and the Correctness of Translations. *J. Theor. Comput. Sci.* **2015**, *577*, 98–124.
21. Gagneur, J.; Klamt, S. Computation of elementary modes: A unifying framework and the new binary approach. *BMC Bioinform.* **2004**, *5*, 175.
22. Madelaine, G.; Lhoussaine, C.; Niehren, J.; Tonello, E. Structural simplification of chemical reaction networks in partial steady states. *Biosystems* **2016**, *149*, 34–49.
23. Schmierer, B.; Tournier, A.L.; Bates, P.A.; Hill, C.S. Mathematical modeling identifies Smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6608–6613.