



**HAL**  
open science

# Simplifications Exactes et Structurelles de Réseaux de Réactions Biologiques

Guillaume Madelaine

► **To cite this version:**

Guillaume Madelaine. Simplifications Exactes et Structurelles de Réseaux de Réactions Biologiques. Informatique [cs]. Université de Lille, 2017. Français. NNT: . tel-01700210

**HAL Id: tel-01700210**

**<https://theses.hal.science/tel-01700210>**

Submitted on 3 Feb 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SIMPLIFICATIONS EXACTES ET STRUCTURELLES DE RÉSEAUX DE RÉACTIONS BIOLOGIQUES

**Guillaume MADELAINE**

**THÈSE**

pour obtenir le titre de Docteur  
de l'École Doctorale Science Pour l'Ingénieur  
**Université de Lille**

**Spécialité : INFORMATIQUE**

Équipe BioComputing, CRIStAL, Université de Lille  
soutenue le 28 février 2017

**Composition du jury :**

<i>Rapporteurs :</i>	Francois FAGES	Inria, Saclay
	Stefan HAAR	Inria, Saclay
<i>Examineurs :</i>	Delphine ROPERS	Inria, Grenoble
	Hélène TOUZET	CNRS, Lille (Inria)
<i>Directeur :</i>	Cédric LHOUSSAINE	Université de Lille
<i>Co-directeur :</i>	Joachim NIEHREN	Inria, Lille

**Université de Lille 1 - Sciences et Technologies**

École Doctorale Sciences Pour l'Ingénieur

Centre de Recherche en Informatique, Signal et Automatique de Lille  
U.F.R. d'I.E.E.A. - Bât. M3 - 59655 VILLENEUVE D'ASCQ CEDEX



## Remerciements

Mes premiers remerciements vont à mes deux directeurs de thèses, Cédric Lhousaine et Joachim Niehren, pour m'avoir guidé durant toutes ces années. Merci pour leurs conseils précieux et pour leurs rigueurs scientifiques, ainsi que pour leurs bienveillances et leurs confiances.

Je remercie ensuite mes rapporteurs, François Fages et Stefan Haar, d'avoir accepté de relire ce document. Je remercie également Delphine Ropers et Hélène Touzet d'avoir accepté de faire partie du jury.

Je tiens à remercier les membres de l'équipe BioComputing, qui m'ont accueilli chaleureusement durant mon doctorat : Émilie pour sa jovialité quotidienne, Mirabelle, qui a encadré mon stage de Master, Quentin et Tom pour leurs participations à l'implémentation de certaines méthodes présentées dans cette thèse, et enfin Cristian, Céline, Kirill et Mathias pour les différentes discussions scientifiques et leurs amabilités.

Je remercie Elisa Tonello et Etienne Farcot pour leur collaboration sur les travaux avec une sémantique déterministe. Je tiens également à remercier particulièrement Michel Petitot pour les très nombreux échanges que nous avons eu pendant cette thèse, et pour les nombreuses idées qu'il a apportées.

J'ai eu la chance durant ce doctorat de pouvoir faire partie des associations de doctorants Tilda et ASPID, et souhaite remercier les doctorants avec qui j'ai partagé cette expérience : Fabien, Antoine, Alexandre, Géraud, Julien et Florian pour Tilda, et Carmelo, Clément, Antoine, Ilkay, Ilyesse, Jason et Théo pour l'ASPID.

Enfin, je remercie ma famille et mes proches pour le soutien qu'ils m'ont apporté et pour leurs confiances.

---

# Table des Matières

---

<b>Table des Matières</b>	<b>3</b>
<b>Index des notations</b>	<b>6</b>
<b>1 Introduction</b>	<b>11</b>
1.1 Contexte . . . . .	11
1.1.1 Biologie des systèmes . . . . .	11
1.1.2 Réseaux de réactions . . . . .	13
1.1.3 Sémantiques des réseaux . . . . .	15
1.1.4 Simplification de réseaux . . . . .	18
1.1.5 Autres approches de modélisation . . . . .	21
1.2 Problématique générale . . . . .	22
1.3 Contributions . . . . .	25
1.3.1 Simplification en non déterministe . . . . .	25
1.3.2 Simplification en déterministe . . . . .	29
1.3.3 Confluence . . . . .	30
1.4 Discussion . . . . .	33
1.5 Plan de la thèse . . . . .	34
<b>2 Préliminaires</b>	<b>37</b>
2.1 Sémantique non déterministe des réseaux . . . . .	38
2.1.1 Réseaux de réactions . . . . .	38
2.1.2 Sémantique opérationnelle . . . . .	41
2.1.3 Exemples biologiques . . . . .	43
2.1.4 Observation . . . . .	47

2.2	Sémantique déterministe . . . . .	50
2.2.1	Réseaux de réactions cinétiques . . . . .	50
2.2.2	Systèmes d'équations . . . . .	54
2.2.3	Sémantique déterministe des réseaux cinétiques . . . . .	56
2.2.4	Exemples biologiques . . . . .	57
2.3	Simplification . . . . .	59
2.3.1	Relation de simplification structurelle . . . . .	59
2.3.2	Simplification correcte . . . . .	60
2.3.3	Confluence . . . . .	61
<b>3</b>	<b>Congruence des attracteurs et simplification non déterministe</b>	<b>65</b>
3.1	Congruence des attracteurs . . . . .	67
3.1.1	Attracteurs . . . . .	67
3.1.2	Convergence et divergence . . . . .	68
3.1.3	Congruence des attracteurs . . . . .	71
3.1.4	Liens avec le graphe des solutions accessibles . . . . .	73
3.2	Ensemble confluent de règles de simplification . . . . .	75
3.2.1	Règles de simplification . . . . .	75
3.2.2	Propriétés . . . . .	77
3.2.3	Application au réseau Tet-On . . . . .	91
3.3	Règles de simplification plus générales . . . . .	94
3.3.1	Règles d'élimination des espèces intermédiaires . . . . .	94
3.3.2	Règles de décomposition . . . . .	102
3.3.3	Règles de symétrie . . . . .	103
3.3.4	Autres règles . . . . .	106
3.3.5	Propriétés . . . . .	108
3.4	Comparaison de modèles : l'exemple de MAPK . . . . .	122
3.5	Discussion . . . . .	125
3.5.1	Changement de fonction d'observation . . . . .	125
3.5.2	Bisimulation . . . . .	127
3.5.3	Épimorphisme de sous-graphe . . . . .	130
3.5.4	Réseaux de Pétri . . . . .	131
3.5.5	Réseaux booléens . . . . .	133
3.6	Conclusion . . . . .	135
<b>4</b>	<b>Simplification préservant la sémantique déterministe</b>	<b>137</b>
4.1	Similarité . . . . .	138
4.2	Règles de simplification . . . . .	141
4.2.1	Simplification d'une espèce intermédiaire . . . . .	141
4.2.2	Combinaison . . . . .	147
4.2.3	Symétrie . . . . .	147

---

4.2.4	Autres règles	150
4.2.5	Propriétés	152
4.3	Application au réseau cinétique Tet-On	159
4.4	Simplification de systèmes d'équations	162
4.5	Discussion	166
4.5.1	Conditions d'équilibre	166
4.5.2	Importance de la structure	167
4.5.3	Modes élémentaires	168
4.5.4	Conditions d'application	169
4.6	Conclusion	170
<b>5</b>	<b>Confluence de l'élimination des espèces intermédiaires</b>	<b>171</b>
5.1	Confluence de la simplification sans cinétique	173
5.1.1	Non confluence de l'élimination des espèces intermédiaires	173
5.1.2	Réseaux de flux	175
5.1.3	Simplification	175
5.1.4	Modes élémentaires	177
5.1.5	Confluence	181
5.2	Confluence structurelle de la simplification avec cinétique	185
5.2.1	Réseaux de flux cinétiques	185
5.2.2	Espèces stables-linéaires	186
5.2.3	Simplification	188
5.2.4	Réseaux linéaires	190
5.2.5	Caractérisation des dépendances	192
5.2.6	Stabilité de <i>LinNets</i>	199
5.2.7	Confluence structurelle	201
5.3	Confluence générale de la simplification avec cinétique	203
5.3.1	Non confluence des expressions cinétiques	203
5.3.2	Critère pour la confluence	203
5.3.3	Confluence	204
5.3.4	Exemple biologique de non confluence des cinétiques	211
5.3.5	Confluence de la simplification de systèmes d'équations	212
5.4	Discussion	215
5.5	Conclusion	218
<b>6</b>	<b>Discussion</b>	<b>219</b>
6.1	Des règles de simplification préservant les deux sémantiques ?	220
6.2	Simplification en stochastique	222
6.3	Conclusion	225
<b>7</b>	<b>Conclusion et perspectives</b>	<b>227</b>

<b>Liste des figures</b>	<b>229</b>
<b>Bibliographie</b> . . . . .	<b>234</b>

## Index des notations

### Réseaux de réactions sans cinétique (Sec. 2.1 et Chapitre 3)

$A, B, X, \dots \in \text{Espèces}$	Espèce moléculaire
$s \in \text{Solutions}$	Solution moléculaire
$r = s_1 \rightarrow s_2 \in \text{Réactions}$	Réaction
$\text{stoic}_r(A) = s_2(A) - s_1(A) \in \mathbb{Z}^n$	Vecteur de stœchiométrie
$\mathbf{R}$	Ensemble de réactions
$N = \mathbf{R} \& s_0 \in \text{Nets}$	Réseau
$\rightarrow \in \text{Nets} \times \text{Nets}$	Sémantique opérationnelle
$\Omega$	Fonction et morphisme d'observation
$o \in \text{Observables}$	Observables
$O$	Ensemble d'observables
$\emptyset$	Observation vide
$\oplus$	Somme sur les observables
$I \subseteq \text{Espèces}$	Molécules internes
$N   N'$	Composition de réseaux
$\{N \rightarrow^*\}$	Ensemble des réseaux accessibles
$\mathcal{G}(N)$	Graphe d'accessibilités
$\underline{\mathcal{G}}_{\text{sol}}(N)$	Graphe des solutions accessibles
$\overline{\mathcal{G}}$	Graphe $\mathcal{G}$ sans les boucles
$\mathcal{A} \in \text{Attr}$	Attracteur
$N \downarrow_O^\Omega$	$N$ peut converger vers $O$
$\uparrow^\Omega$	$N$ peut diverger
$\simeq_\Omega$	Équivalence des attracteurs
$\sim_{\Omega, I}$	Congruence des attracteurs
$\Rightarrow_{\Omega, I}$	Simplification



## Réseaux de réactions cinétiques (Sec. 2.2 et Chapitre 4)

$x_A \in Vars$	Variable pour la concentration de A
$x_A^0 \in Vars$	Variable pour la concentration initiale de A
$k \in Param$	Paramètre
$e \in Expr$	Expression
$\beta : Param \rightarrow \mathbb{R}_{>0}$	Interprétation des paramètres
$\alpha : Vars \rightarrow (\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0})$	Interprétation des variables
$\llbracket e \rrbracket_\alpha \in \mathbb{R} \cup \{undef\}$	Interprétation d'une expression
$C \in Contraintes$	Contrainte
$\llbracket C \rrbracket_\alpha \in \mathbb{B} \cup \{undef\}$	Interprétation d'une contrainte
$r = s_1 \rightarrow s_2 ; e$	Réaction cinétique
$R$	Ensemble de réactions cinétiques
$M = R \& C \in KinNets$	Réseau cinétique
$\dot{x} = e$	Équation différentielle ordinaire
$E \in Systemes$	Système d'équations
$\Omega$	Fonction d'observation
$ObsAtoms$	Ensemble des observations atomiques
$\Omega(\alpha)$	Interprétation observable
$sol_\Omega(E)$	Ensemble des solutions observables de E
$\equiv$	Similarité
$\simeq_\Omega$	Équivalence de systèmes d'équations
$\simeq_\Omega$	Équivalence de réseaux cinétiques
$\sim_{\Omega, I}$	Congruence de réseaux cinétiques
$\Rightarrow_{\Omega, I}$	Simplification de réseaux cinétiques

## Réseaux de flux (Chapitre 5)

$\mathbf{r} = (r_1, \dots, r_m)$	Base/vecteur de réactions (sans cinétique)
$u \in \mathbb{Z}^m$	Flux
$u\mathbf{r}$	Réaction du flux $u$ dans la base $\mathbf{r}$
$U$	Réseau de flux
$v = u; e$	Flux cinétique
$v\mathbf{r} = u\mathbf{r}; e$	Réaction cinétique du flux $u$ dans la base $\mathbf{r}$
$W = V\&C$	Réseau de flux cinétiques, avec contraintes
$M = \{v\mathbf{r} \mid v \in V\}\&C$	Réseau de réactions correspondant
$\mathcal{J} \subseteq \text{Espèces}$	Ensemble des espèces intermédiaires
$S$	Matrice de stœchiométrie d'une base $\mathbf{r}$
$S_{\mathcal{J}}$	Projection de $S$ sur $\mathcal{J}$
$\text{Reac}_{\mathcal{J}}(v\mathbf{r})$	Espèces intermédiaires consommées par $v\mathbf{r}$
$\text{Prod}_{\mathcal{J}}(v\mathbf{r})$	Espèces intermédiaires produites par $v\mathbf{r}$
$\text{LinNets}$	Ensemble de réseaux linéaires, stable pour la simplification
$\tilde{u}$	Chemin ou chemin de flux
$\equiv$	Similarité
$\equiv^{struc}$	Similarité structurelle
$\simeq_{\Omega}$	Équivalence de réseaux de flux
$\sim_{\Omega, \mathcal{I}}$	Congruence de réseaux de flux
$\Rightarrow_{\Omega, \mathcal{I}}$	Simplification de réseaux de flux



# CHAPITRE 1

---

## Introduction

---

### 1.1 Contexte

#### 1.1.1 Biologie des systèmes

La *biologie des systèmes* ([Kitano, 2002](#)) est un domaine inter-disciplinaire, avec pour objectif de comprendre et d'analyser des systèmes biologiques à l'aide de modèles. Cette approche a émergée suite à l'explosion des données expérimentales en biologie, ainsi que l'amélioration de la qualité de ces données, conséquence de l'évolution des technologies et de la recherche, et notamment des méthodes de séquençage de l'ADN. La compréhension du fonctionnement d'un système biologique complexe (réseau métabolique, cycle circadien, etc.) ne peut pas se faire en analysant uniquement ses composants (molécules, ADN, etc.) de façon individuelle. En effet, sa dynamique dépend des *interactions* entre les différents composants, et la biologie des systèmes adopte donc une vue globale en considérant le comportement du système biologique dans son ensemble.

L'objectif de la biologie des systèmes est alors de *comprendre* le comportement des systèmes, à travers des *modèles*. Les modèles sont *construits* à partir de *données expérimentales*, avec des méthodes mathématiques et informatiques. Plusieurs étapes de validation expérimentale et de reconstruction du modèle sont en général nécessaire pour obtenir un modèle correct, c'est-à-dire dont les résultats ne sont pas contredits par les expériences ([Popper, 2005](#)). Le type de modèle et son niveau de détails dépendent alors de la quantité et la qualité des données expérimentales. Parfois, les seules données disponibles sont des corrélations et des

dépendances entre les différents composants du système. Il n'est alors possible que de créer des réseaux de relations, comme par exemple dans (Butte & Kohane, 2000), ou des réseaux d'influence entre espèces, comme dans (A. Rao, Hero III, & Engel, 2007). Lorsque plus de données sont disponibles, il est possible de construire des réseaux de réactions, représentant des transformations entre espèces. C'est par exemple le cas du réseau métabolique d'algues proposé dans (Prigent et al., 2014), ou de la bactérie *Bacillus Subtilis* proposé dans (Coutte et al., 2015). Ces deux modèles ne possèdent pas d'informations sur les cinétiques des réactions, mais il est toujours possible de faire des analyses qualitatives, ou dans un état d'équilibre. Lorsque des données sur la cinétique sont disponibles, on peut assigner à chaque réaction une expression cinétique représentant sa vitesse, comme par exemple dans les modèles du cycle cellulaire (Qu, MacLellan, & Weiss, 2003), ou de l'horloge circadienne (Leloup & Goldbeter, 2003). Néanmoins, les valeurs des paramètres cinétiques ne sont en général connues que pour des petits modèles précis, comme par exemple pour le réseau Tet-On (Huang, Moya, Jayaraman, & Hahn, 2010) ou le réseau MAPK (Markevich, Hoek, & Kholodenko, 2004), tandis que les gros modèles restent incomplets.

Lorsque les modèles sont suffisamment complets, ils peuvent être *simulés*, c'est-à-dire qu'il est possible de reproduire leurs évolutions dans le temps sur ordinateur. Le programme de simulation peut être déduit automatiquement du modèle, par exemple avec l'algorithme de Gillespie (1977) ou par intégration numérique d'équations différentielles ordinaires. Cela permet alors de *prédire* de nouveaux comportements, par exemple en réponse à des modifications de l'environnement biologique. De nouvelles expériences peuvent alors permettre de *valider* (Heiner, Koch, & Will, 2004) ou au contraire *réfuter* (Anderson & Papachristodoulou, 2009) le modèle. En général, il n'est pas possible d'obtenir un modèle imitant à 100% les données obtenues expérimentalement, celles-ci étant souvent très bruitées. On peut néanmoins valider un modèle s'il est suffisamment proche de la réalité, et s'il permet bien de simuler les propriétés importantes du système. Mais même un modèle valide ne représentera souvent qu'une partie des composantes ou interactions du système, et sera donc incomplet. Lorsque le modèle est suffisamment petit et le nombre de paramètres est faible, leurs valeurs peuvent être *estimées* (Ashyraliyev, Fomekong-Nanfack, Kaandorp, & Blom, 2009) pour que le comportement du modèle concorde avec les données expérimentales. L'*analyse formelle* du modèle permet alors de mieux comprendre le fonctionnement du système biologique, et de découvrir des propriétés non triviales du système, comme par exemple la présence d'oscillations. Un modèle peut servir à *prédire in silico* des comportements possibles du système, non observés expérimentalement. On peut par exemple chercher à savoir quel serait l'effet d'une modification du système ou de son environnement, par exemple en injectant (de façon ponctuelle ou continue) une nouvelle protéine, ou en désactivant un gène (Burgard, Pharkya, &

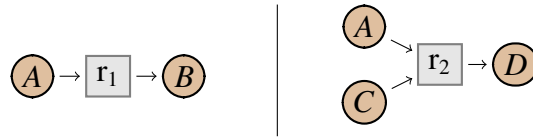


Fig 1.1: Exemple de réactions chimiques. La réaction  $r_1 = A \rightarrow B$  transforme une molécule de l'espèce  $A$  en une molécule de l'espèce  $B$ . La réaction  $r_2 = A + C \rightarrow D$  transforme une molécule de  $A$  et une molécule de  $C$  en une molécule de  $D$ .

(Maranas, 2003; Coutte et al., 2015). Un modèle peut aussi permettre à essayer de *contrôler* un système, en l'observant et en pouvant modifier en temps réel son environnement ( par exemple à l'aide d'un dispositif micro-fluidique), afin de le faire atteindre un comportement cible (Uhlendorf et al., 2012). Comme pour les modèles informatiques, on peut utiliser des méthodes de *vérification* (Queille & Sifakis, 1982), par exemple pour vérifier qu'un système n'atteint jamais un comportement critique (comme la mort de la cellule), pour savoir si la présence de telle protéine est nécessaire pour l'activation de tel gène, ou pour tester des propriétés de logique temporelle (voir par exemple *BioCham* (Fages, Soliman, & Chabrier-Rivier, 2004)). On peut également étudier ses *propriétés dynamiques*, pour savoir s'il peut se stabiliser au cours du temps, s'il diverge en produisant une infinité de molécules, ou s'il oscille entre différents états. En *biologie synthétique*, des modèles sont utilisés pour construire artificiellement de nouveaux systèmes biologiques (Andrianantoandro, Basu, Karig, & Weiss, 2006; Purnick & Weiss, 2009).

### 1.1.2 Réseaux de réactions

Dans cette thèse, nous modélisons les systèmes biologiques à l'aide de *réseaux de réactions chimiques*, pour un ensemble d'espèces chimiques (abstraites) donné.

Une *réaction chimique* (sans cinétique) décrit une transformation d'un ensemble d'espèces chimiques en un autre ensemble. Par exemple, la réaction  $G_i \rightarrow G_a$  transforme un gène inactif  $G_i$  en un gène actif  $G_a$ , tandis que la réaction  $G_a \rightarrow G_a + P$  décrit la production d'une protéine  $P$  lorsque le gène est actif. Les composants chimiques sont ici abstraits : on ne considère pas par exemple leurs structures atomiques (au contraire par exemple des réactions de type  $2H + O \rightarrow H_2O$ ). Ainsi, une espèce  $A$  peut représenter un complexe de protéines, un gène, etc. Les réactions peuvent représenter l'activation d'un gène, la liaison de deux molécules, la dégradation d'une protéine, le déplacement d'une molécule entre deux compartiments, etc.

Une réaction chimique peut alors être représentée graphiquement. Par exem-

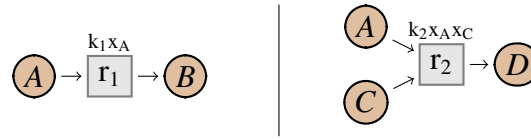


Fig 1.2: Exemple de réactions chimiques avec cinétiques. La vitesse  $k_1 x_A$  de la réaction  $r_1 = A \rightarrow B$ ;  $k_1 x_A$  dépend d'un paramètre  $k_1$  et de la quantité  $x_A$  de l'espèce  $A$ . La vitesse de la réaction  $r_2 = A + C \rightarrow D$ ;  $k_2 x_A x_C$  est  $k_2 x_A x_C$ .

ple, les réactions  $r_1 = A \rightarrow B$  et  $r_2 = A + C \rightarrow D$  sont représentées dans la Fig. 1.1. La première réaction représente la transformation d'une molécule d'une espèce  $A$  en une autre espèce  $B$ , tandis que la deuxième réaction décrit la liaison d'une molécule de  $A$  avec une molécule de  $C$ , formant un complexe noté  $D$ .

Une *réaction chimique avec cinétique* est une paire composée d'une réaction sans cinétique et d'une *expression cinétique*. Une expression est un terme composé de variables et de paramètres, combinés grâce à des opérateurs mathématiques, et représente la vitesse de la réaction. L'expression peut, selon la sémantique, être interprétée de façon déterministe ou stochastique. Par exemple, dans la Fig. 1.2, la vitesse  $k_1 x_A$  de la réaction  $r_1 = A \rightarrow B$ ;  $k_1 x_A$  est égale au produit du paramètre  $k_1$  et de la quantité  $x_A$  de l'espèce  $A$ . Selon la sémantique, cette quantité peut être discrète ou continue.

Un *réseau de réactions* est alors un ensemble de réactions chimiques, qui décrit les différentes interactions entre les espèces d'un système biologique. Par exemple, le réseau de réactions cinétiques du système Tet-On (Gossen & Bujard, 1992; Gossen et al., 1995; Huang et al., 2010) est représenté dans la Fig. 1.3. Les flèches grisées en pointillées représentent les modificateurs d'une réaction, c'est-à-dire des espèces qui influent sur la vitesse de la réaction, mais dont la quantité n'est pas modifiée par cette réaction. Par exemple, dans la réaction  $r_1 = \emptyset \rightarrow Dox_i$ ;  $k_1 x_{Dox}$ , qui produit une molécule de  $Dox_i$ , l'espèce  $Dox$  est un modificateur. On étudiera ce réseau en détails plus loin. Brièvement, il modélise, à travers une chaîne de réactions, la production d'une protéine fluorescente verte ( $GFP_a$ ) en présence de doxycycline ( $Dox$ ) à l'extérieur d'une cellule. Les réseaux de réactions ont alors trois sémantiques classiques, que l'on décrira plus en détails plus loin : la sémantique non déterministe, la sémantique déterministe et la sémantique stochastique.

Un réseau de réactions, sans cinétique, est en fait équivalent à un *réseau de Pétri* non marqué (Petri, 1966; Murata, 1989; Silva, 2013). Chaque place correspond alors à une espèce chimique et chaque transition à une réaction. Un réseau de Pétri marqué est équivalent à un réseau de réactions doté d'une solution initiale. Les réseaux de Pétri sont très utilisés et étudiés en biologie des systèmes

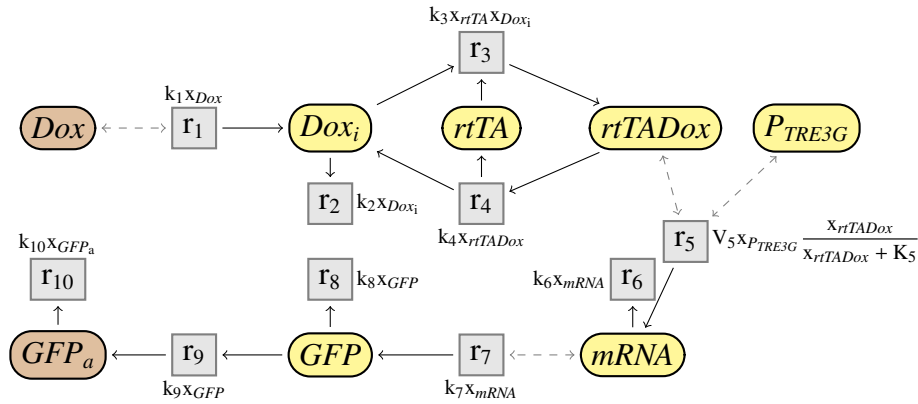


Fig 1.3: Réseau de réactions cinétiques du système Tet-On.

(Reddy, Mavrouniotis, & Liebman, 1993; Chaouiya, 2007; Heiner, Gilbert, & Donaldson, 2008). En ajoutant des cinétiques, les réseaux de Pétri continus (Alla & David, 1998) et stochastiques (Bause & Kritzinger, 2002) permettent également de représenter les sémantiques déterministe et stochastique des réseaux de réactions.

Plusieurs outils permettent alors de modéliser les systèmes de réactions sous forme de réseaux, et de les analyser. *BioCham* (Fages et al., 2004) propose par exemple une syntaxe pour les réseaux de réactions avec cinétiques. Il propose également un ensemble d'outils, permettant de simuler les réseaux, ou bien de les analyser, par exemple en utilisant des méthodes de vérification pour vérifier des propriétés de logiques temporelles. L'outil COPASI (Hoops et al., 2006) permet également de modéliser des réseaux de réactions, et de les simuler, avec une sémantique déterministe, stochastique ou hybride.

### 1.1.3 Sémantiques des réseaux

Plusieurs sémantiques sont alors possibles pour les réseaux de réactions, représentant différents types de dynamiques possibles, selon qu'on considère le temps de façon continue, ou bien juste comme une succession de transformation, et selon si les quantités de matières sont vues de façon discrète ou continue (voir par exemple Érdi & Tóth, 1989).

Pour des réseaux sans cinétique, une possibilité est d'adopter une *sémantique non déterministe* pour les réseaux de réactions, similairement à la sémantique classique des réseaux de Pétri (1966). La sémantique opérationnelle est une relation entre des multi-ensembles de molécules. Elle permet de transformer des



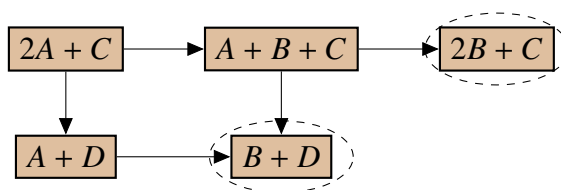


Fig 1.4: Graphe d'accessibilité pour les réactions de la Fig. 1.1, avec initialement deux molécules de A et une molécule de C. Les cercles pointillés représentent les états terminaux.

molécules en d'autres molécules en appliquant l'une des réactions du réseau. Plusieurs réactions peuvent être applicables en même temps. Cette sémantique considère alors tous les choix possibles, de façon non déterministe. Il n'y a pas vraiment de notion de temps ici, mais plutôt une succession d'états. Cette sémantique permet de construire le *graphe d'accessibilité* du réseau, représentant l'ensemble de tous les comportements possibles, et non pas juste le comportement moyen comme dans la sémantique déterministe. Par exemple, le graphe d'accessibilité du réseau composé des deux réactions de la Fig. 1.1, en prenant initialement deux molécules de A et une molécule de C, est représenté dans la Fig. 1.4.

La *sémantique déterministe* d'un réseau de réactions avec cinétiques décrit l'évolution au cours du temps du nombre moyen de molécules dans un volume fixé. La quantité de matières est représentée ici de façon continue, par une *concentration*, un nombre réel correspondant à une quantité de molécules par unité de volume. Cette approche n'est valable que sous certaines conditions, en particulier la quantité de molécules doit être suffisamment grande, répartie de façon homogène dans le volume, etc. Dans ce cas, le comportement déterministe peut être vu comme la moyenne des comportements stochastiques. Pour une concentration initiale donnée, le comportement du réseau est alors complètement déterministe.

Le comportement est alors décrit par un ensemble d'*équations différentielles ordinaires* (voir par exemple [Heinrich & Schuster, 2012](#)). Par exemple, le système d'équations correspondant aux réactions de la Fig. 1.1 est donné dans la Fig. 1.5. Lorsque les valeurs des paramètres cinétiques et des concentrations initiales sont connues, ces systèmes peuvent être *simulés* de façon numérique (Fig. 1.5). Dans de très rares cas, lorsque les systèmes sont suffisamment petits et simples, on peut les *résoudre* de façon symbolique (voir par exemple [Cornish-Bowden, 2012](#)).

Lorsque les données cinétiques ne sont pas connues (ou que partiellement), ce qui est souvent le cas pour les grands systèmes, il est alors possible de n'étudier que le comportement *stable* du système, c'est-à-dire lorsque les concentrations

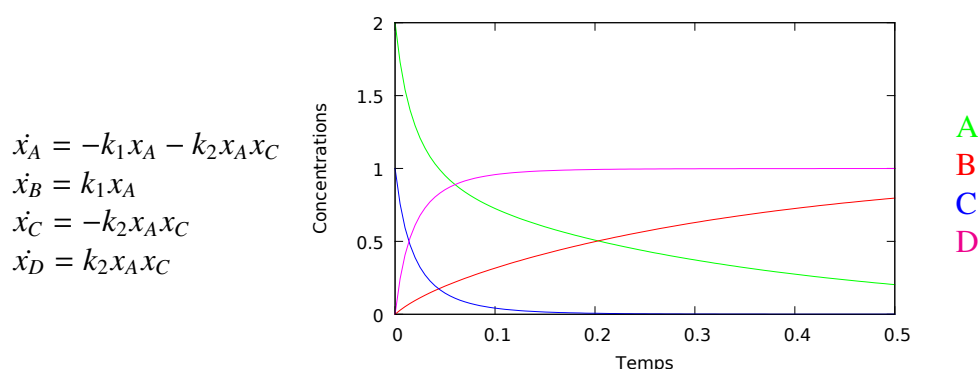


Fig 1.5: Exemple de système d'équations différentielles et de l'évolution des concentrations des réactions de la Fig. 1.1, pour des valeurs de paramètres et des concentrations initiales fixées. La concentration de A est en vert, celle de B en rouge, celle de C en bleu et celle de D en magenta.

des différentes espèces ont atteint un équilibre et restent constantes au cours du temps. De nombreuses méthodes d'analyse supposent que le système est dans un état d'équilibre, ou proche de cet état. C'est par exemple le cas de techniques d'analyse de flux (Orth, Thiele, & Palsson, 2010) ou de simplification de systèmes d'équations (Bodenstein, 1913).

La *sémantique stochastique* décrit quand à elle la dynamique probabiliste des réseaux, en temps continu. Pour tout temps  $t \in \mathbb{R}_{\geq 0}$ , elle donne une distribution de probabilité entre différents états (discrets). La stochasticité est un élément central de nombreux systèmes biologiques, notamment lorsque le nombre de molécules de certaines espèces est faible (par exemple pour les gènes). Elle nécessite de connaître les cinétiques des réactions.

La dynamique stochastique des réseaux de réactions peut être décrite à l'aide de *chaîne de Markov à temps continu* (Norris, 1998). Elle peut également être représentée par l'*équation maîtresse chimique* (D. T. Gillespie, 1977), un système d'équations différentielles sur la probabilité d'être dans un certain état au temps  $t$ . Ce système ne peut malheureusement pas être résolu dans la plupart des cas, le nombre d'équations différentielles étant infini. Il peut néanmoins être simulé, grâce à l'algorithme de simulation stochastique, aussi appelé algorithme de Gillespie (1977). La Fig. 1.6 donne un exemple d'une simulation stochastique pour les réactions de la Fig. 1.2. Une autre approche, lorsque les réseaux sont trop grands pour être analysés ou simulés directement, consiste à approximer l'équation maîtresse en utilisant des méthodes de clôture de moments (C. S. Gillespie, 2009).

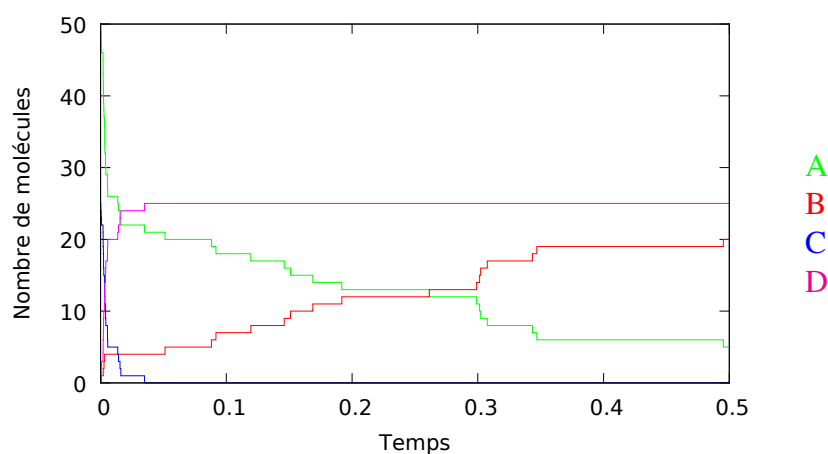


Fig 1.6: Exemple d'une simulation stochastique des réactions de la Fig. 1.2, avec initialement 50 molécules de A et 25 molécules de C. Le nombre de molécules de A est en vert, le nombre de B en rouge, celui de C en bleu et celui de D en magenta.

### 1.1.4 Simplification de réseaux

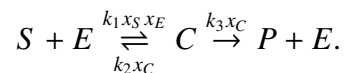
Les réseaux deviennent aujourd'hui de plus en plus grands, en cherchant à modéliser des systèmes biologiques complets avec suffisamment de détails. Par exemple, la base de données *SubtiWiki* (Mäder, Schmeisky, Flórez, & Stülke, 2011) regroupe les informations sur les gènes et les protéines de la bactérie *Bacillus Subtilis*. Ils proposent également un modèle des interactions protéines, avec plus de 800 protéines et 1800 réactions. Le projet *Recon 2* (Thiele et al., 2013) tente quant à lui de construire un modèle complet du métabolisme humain, avec plus de 5000 métabolites et 7000 réactions. Les modèles avec des informations cinétiques sont en général plus petits, mais deviennent eux aussi de plus en plus gros avec l'amélioration des techniques expérimentales et de la qualité des données biologiques obtenues. Par exemple, le réseau cinétique du système complet de la signalisation MAPK proposé dans (Schoeberl, Eichler-Jonsson, Gilles, & Müller, 2002) contient environ 100 espèces et 250 réactions.

L'objectif de la *simplification* de réseaux est de réduire la taille des gros modèles. Avoir des *petits réseaux* présentent des avantages évidents. Comme expliqué précédemment, la résolution (symbolique) des systèmes d'équations différentielles ou de l'équation maîtresse ne peut se faire que pour des réseaux simples. Lorsque les informations cinétiques sont connues, un réseau plus petit permet de faire des simulations plus rapides. Dans le cas contraire, il est plus facile d'estimer les paramètres d'un petit modèle, or ces données sont souvent difficiles

à obtenir expérimentalement. Il est plus facile de prédire le comportement de petits systèmes, en particulier lorsqu'il manque des informations cinétiques. Un petit modèle est évidemment plus simple à comprendre et à analyser, et il est plus facile d'identifier les mécanismes clés responsables de ses comportements. Il est également plus facile à contrôler.

Les biologistes, lors de la construction de modèles, procèdent souvent à des simplifications *intuitives*, en remplaçant un sous-système par une unique "méta-réaction". C'est par exemple le cas de la production de protéines par un gène, qui est souvent représentée par une unique réaction. En réalité, elle décrit un système complexe : le gène doit être activé, puis la polymérase doit trouver le gène. Elle procède alors à sa transcription, qui produit de l'ARN messenger. Celui-ci doit quitter le noyau et rejoindre le cytoplasme (pour les cellules eucaryotes), puis être traduit en protéine par les ribosomes. Ces simplifications entraînent souvent la perte des informations sur les fonctions cinétiques. Et même lorsque les cinétiques sont conservées, ces simplifications ne garantissent pas que le comportement du réseau simplifié soit le même que celui du réseau initial, et peuvent être une source d'erreurs. Par exemple, il est montré dans (Del Vecchio, 2007) qu'un modèle d'expression génétique avec ARN messenger n'a pas les mêmes comportements qu'un modèle où cette étape est simplifiée.

L'exemple classique de simplification de réseaux calculant la cinétique du réseau simplifié est la simplification des transformations enzymatiques proposée par Michaelis et Menten (1913). Le système biologique transforme un substrat  $S$  en un produit  $P$  en présence d'une enzyme  $E$ . Le réseau initial (à gauche dans la Fig. 1.7) est souvent décrit dans la forme suivante :



Le système est composé de trois réactions avec des cinétiques de type loi d'action de masse. Premièrement, le substrat  $S$  se lie à l'enzyme  $E$ , formant un complexe  $C$ . Cette réaction est réversible : le complexe peut se décomposer en  $S$  et  $E$ . Il peut aussi produire le produit  $P$  en relâchant l'enzyme. Toutes ces réactions ont une cinétique respectant la loi d'action de masse. Les espèces  $E$  et  $C$  (en jaune dans la figure) sont des espèces considérées comme *intermédiaires*, c'est-à-dire des espèces qui agissent sur le comportement du système, mais qui ne sont que des étapes et ne nous intéressent pas particulièrement. Michaelis et Menten proposent alors de simplifier le réseau en supprimant ces deux espèces intermédiaires. Ils obtiennent alors un réseau avec une unique réaction, transformant directement  $S$  en  $P$ , avec une cinétique particulière, appelée cinétique de Michaelis-Menten. Ils montrent que lorsque le complexe  $C$  est proche de l'équilibre, le comportement du réseau simplifié est alors une approximation du comportement du réseau initial.

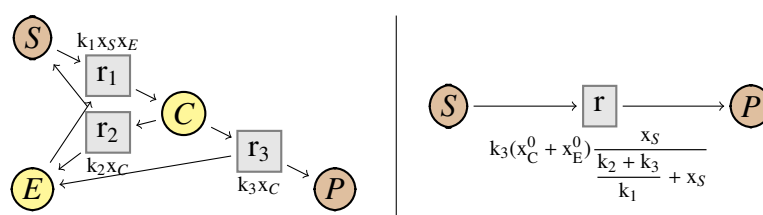


Fig 1.7: Réseau initial (gauche) et simplifié (droite) pour Michaelis-Menten.

Notons que si le complexe est exactement à l'équilibre, alors le comportement est préservé de façon exacte.

À l'exception de l'exemple de Michaelis-Menten, dans la littérature, la majorité des simplifications existantes réduisent directement les *équations différentielles*, (voir par exemple [Okino & Mavrouniotis, 1998](#) pour un résumé des différentes approches), en se servant parfois de la structure du réseau initial ([King & Altman, 1956](#)). Ces méthodes ne calculent que les équations simplifiées, et pas la structure du réseau simplifié.

À l'inverse, les simplifications entièrement basées sur la structure, sans cinétique, ne préservent en général pas de notion de sémantique. C'est par exemple le cas la simplification proposée dans ([Gay, Soliman, & Fages, 2010](#)), basée sur les épimorphismes de sous-graphe ([Gay, Fages, Martinez, Soliman, & Solnon, 2014](#)) dans des réseaux sans information cinétique. Leur simplification cherche à relier entre eux des modèles existants, mais ne s'occupe pas des questions de sémantique, et donc il n'y a pas de notion d'équivalence entre le réseau initial et celui simplifié. Toujours sans cinétique, les simplifications dans les réseaux de Pétri (voir par exemple les travaux de Berthelot et al. ([1976](#); [1980](#); [1985](#))) permettent par exemple d'éliminer des places (espèces) intermédiaires dans les réseaux. Mais ces simplifications visent à préserver des propriétés spécifiques des réseaux de Pétri, comme la vivacité ou l'existence d'une borne sur le nombre total de molécules, et pas une sémantique non déterministe plus générale.

Avec des réseaux de réactions avec cinétiques et sous condition d'une sémantique déterministe, Radulescu et al. ([2008](#)) proposent une simplification qui calcule séparément la structure simplifiée, en se basant sur les modes élémentaires ([Schuster & Hilgetag, 1994](#)), et le système simplifié d'équations. Lorsque c'est possible, ils assignent alors à chaque réaction simplifiée un taux cinétique afin de vérifier le système simplifié. Pour simplifier le système d'équations, ils utilisent des approximations sur les cinétiques, sans donner de résultats de correction.

### 1.1.5 Autres approches de modélisation

D'autres *approches de modélisation* peuvent être utilisés pour représenter des systèmes biologiques, en utilisant des méthodes mathématiques et informatiques.

Par exemple, les *réseaux booléens* (Kauffman, 1969; Thomas, 1973) sont des abstractions, dans lesquelles une concentration peut être soit égale à zéro (l'espèce est inactive), soit différente de zéro (l'espèce est active). Le temps est ici discret. Ils sont en général utilisés pour représenter des réseaux de gènes, et permettent notamment de caractériser la présence d'oscillations.

Des *modèles hybrides* (Alur et al., 2001) peuvent être utilisés pour représenter certaines parties du système de façon déterministe (lorsque le nombre de molécules est élevé), et d'autres parties de façon stochastique (pour un faible nombre de molécules).

Lorsque l'on veut modéliser des molécules complexes, sous forme de graphe, comme par exemple des séquences de molécules (polymérisation, ADN, etc.), il est possible d'utiliser des langages à base de règles (Danos & Laneve, 2004; Blinov, Faeder, Goldstein, & Hlavacek, 2004; John, Lhousseine, Niehren, & Versari, 2011). Une règle est alors une réaction paramétrique, une méta-réaction, représentant de façon compacte un ensemble de réactions.

Une autre approche consiste à s'inspirer des langages de programmations, les systèmes biologiques étant concurrents par nature. On peut par exemple utiliser des langages dérivés du  $\pi$ -calcul (Regev, Panina, Silverman, Cardelli, & Shapiro, 2004; Kuttler, Lhousseine, & Niehren, 2007), du join-calcul (Fournet & Gonthier, 1996) ou des algèbres de processus (Ciocchetta & Hillston, 2008). Cette approche permet d'utiliser les notions de sémantiques et les méthodes d'analyse développées pour les langages de programmation (Mitchell, 1996; Winskel, 1993; Milner, 1989).

## 1.2 Problématique générale

La problématique générale est alors de trouver des *algorithmes de simplification* pour les réseaux de réactions, qui préservent leurs sémantiques. En particulier, la simplification doit être capable d'éliminer des espèces intermédiaires. Il y a alors 5 défis à surmonter : avoir une simplification qui soit à la fois *riche*, *correcte*, *structurelle*, *contextuelle* et *confluente*.

Tout d'abord, on veut que la simplification soit *riche*, c'est-à-dire qu'elle puisse s'appliquer à un maximum de réseaux, et surtout à des réseaux venant des applications biologiques. Elle doit pouvoir supprimer le plus de détails possibles, sans modifier le comportement global du réseau. En particulier, la simplification doit être capable de supprimer des espèces intermédiaires, mais aussi d'éliminer des symétries, etc.

Par exemple, le réseau de réactions représenté graphiquement dans la Fig. 1.8 (à gauche) transforme une molécule de  $A$  en une molécule de  $X$ , puis soit en  $Y_1$  soit en  $Y_2$ , et ces molécules se transforment alors en  $B$ . On voit directement que globalement, le réseau va juste transformer  $A$  en  $B$ . On veut alors le simplifier pour obtenir le réseau de droite, qui transforme directement  $A$  en  $B$ .

On veut également que la simplification soit *correcte*, c'est-à-dire que le comportement du réseau simplifié soit le même que celui du réseau initial. Cette propriété est indispensable, puisque l'on souhaite pouvoir analyser ou simuler le réseau simplifié et en déduire des informations sur le réseau initial. Il faut donc que ces deux réseaux soient *équivalents*. La notion d'équivalence va dépendre de la sémantique considérée, et devra représenter les comportements qui nous intéressent. Avec une sémantique déterministe, deux réseaux vont par exemple être équivalents si leurs systèmes d'équations différentielles ont les mêmes solutions. Pour le cas non déterministe, l'équivalence est plus difficile à définir. Elle doit pouvoir représenter le comportement final et régulier des réseaux. Chaque règle de simplification doit alors être correcte pour l'équivalence, c'est-à-dire préserver la sémantique.

On veut ensuite que la simplification d'un réseau de réactions soit encore un réseau de réactions, avec une *structure*, et pas juste un ensemble d'équations différentielles. En effet, connaître la structure des réseaux de réactions apporte plusieurs avantages, pour leurs compréhensions et pour leurs simplifications.

Premièrement, ils peuvent être représentés *graphiquement*, et sont alors plus simple à comprendre qu'un ensemble d'équations. On peut alors facilement et rapidement identifier le comportement global, essentiel, du réseau.

De plus, il est facile de dériver à la fois la sémantique non déterministe,

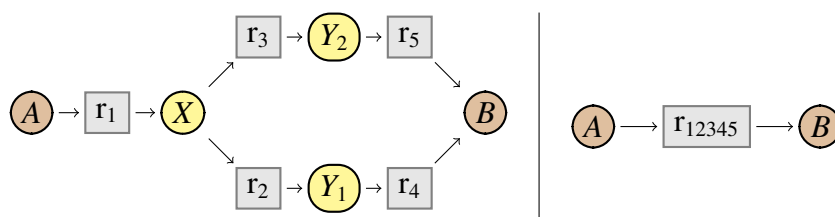


Fig 1.8: Exemple de simplification du réseau de gauche pour obtenir le réseau de droite. Les espèces  $X$ ,  $Y_1$  et  $Y_2$  sont des espèces intermédiaires, que l'on souhaite supprimer.

la sémantique déterministe et la sémantique stochastique à partir d'un réseau de réactions et de sa structure. Les réseaux constituent donc un bon choix de modélisation. À l'inverse, retrouver un réseau de réactions à partir de son système d'équations différentielles est un problème difficile (Fages, Gay, & Soliman, 2015). En effet, plusieurs réseaux peuvent produire le même système, et il est donc délicat de trouver le "meilleur" réseau, le plus cohérent, pour un système donné.

Une autre difficulté est d'obtenir une simplification *contextuelle*. L'idée ici est de pouvoir prendre un sous-réseau, un *module* d'un réseau plus grand, et de pouvoir le remplacer par un autre module équivalent. Cette opération ne doit pas modifier le comportement global du grand réseau, donc le grand réseau après ce remplacement doit être équivalent au grand réseau initial. Cette approche contextuelle nécessite de connaître la structure du réseau de réactions, elle est difficilement envisageable avec des équations différentielles.

Une telle approche contextuelle permet de simplifier un réseau facilement, de façon *locale*. Elle permet d'obtenir des règles de simplification simples, qui se concentrent sur le module à simplifier. L'exemple classique ici est la simplification proposée par Michaelis et Menten (1913) : dans un grand réseau, on remplace une transformation enzymatique composée de trois réactions par une unique réaction (voir Fig. 1.7 plus haut). La simplification est alors valable quelque soit le contexte, à condition que celui-ci n'interagisse pas avec les espèces  $E$  et  $C$ , et respecte les différentes conditions d'équilibre nécessaires à la simplification.

Une autre problématique importante est d'avoir une simplification *confluente*, c'est-à-dire que quelque soit l'ordre d'application des règles de simplification, on obtient toujours le même réseau simplifié. Ainsi, le réseau complètement simplifié est *unique*. Quelque soit l'ordre de la simplification, on est alors sûr de toujours obtenir le plus petit possible. Il n'y a alors pas d'ordre meilleur que les autres



(d'un point de vue du résultat de la simplification), et il n'est donc pas nécessaire de tester toutes les possibilités. L'ordre d'application des règles peut alors être choisi de façon à *optimiser* le temps de calcul de la simplification. De plus, avoir un unique réseau complètement simplifié permet de plus facilement utiliser la simplification pour *comparer* différents modèles d'un même système biologique.

## 1.3 Contributions

Dans cette thèse, nous cherchons à simplifier les systèmes à réactions chimiques en biologie systémique en adoptant l'approche suivante. Dans un premier temps, nous négligeons les cinétiques et développons une approche purement structurelle, pour une sémantique non déterministe. On essaye alors dans un deuxième temps d'étendre les règles de simplifications en ajoutant les cinétiques.

### 1.3.1 Simplification en non déterministe

Dans le cadre sans cinétique avec une sémantique opérationnelle non déterministe, le problème est double. Il faut en effet trouver d'une part une *simplification structurelle* des réseaux de réactions. Et d'autre part il faut définir, à partir de la sémantique opérationnelle non déterministe, une nouvelle *sémantique contextuelle*, qui soit pertinente biologiquement, et telle que les règles de simplification soient *correctes* relativement à cette sémantique contextuelle.

Pour illustrer les problématiques dans le cadre non déterministe, on considère le système Tet-On (Gossen & Bujard, 1992; Gossen et al., 1995; Huang et al., 2010), représenté dans la Fig. 1.9. Dans ce réseau, la présence de doxycycline (*Dox*) à l'extérieur d'une cellule enclenche une chaîne de réactions, qui aboutie à la production d'une protéine fluorescente verte (*GFP<sub>a</sub>*). Ici, le contexte ne peut interagir qu'avec la doxycycline à l'extérieur de la cellule, et pas avec les autres espèces dans la cellule. On observe alors la protéine fluorescente dans la cellule. Intuitivement, le réseau semble donc équivalent au réseau de la Fig. 1.10 où *GFP<sub>a</sub>* est produite directement en présence de *Dox*.

On veut alors avoir une simplification permettant de transformer le réseau initial de Fig. 1.9 pour obtenir le réseau simple de Fig. 1.10. Et on veut également pouvoir justifier formellement le fait que ces deux réseaux se comportent de la même façon. C'est-à-dire que pour tout contexte n'interagissant qu'avec *Dox*, le comportement de l'espèce observable *GFP<sub>a</sub>* est le même dans les deux réseaux. Les autres espèces sont alors toutes considérées comme des espèces intermédiaires.

Nous proposons tout d'abord un ensemble de règles de *simplifications* de réseaux de réactions. Ces règles *structurelles* permettent de simplifier de façon locale un réseau. On peut partitionner ces règles en différentes catégories.

Un premier ensemble de règles vise à éliminer une *espèce intermédiaire*. Le principe de ces règles, représenté dans la Fig. 1.11, est de combiner deux-à-deux toute réaction produisant l'espèce intermédiaire avec toute réaction la consommant. Par exemple, on peut appliquer ce type de règles pour supprimer l'espèce *GFP* dans le réseau de Tet-On de la Fig. 1.9. On obtient alors le réseau par-

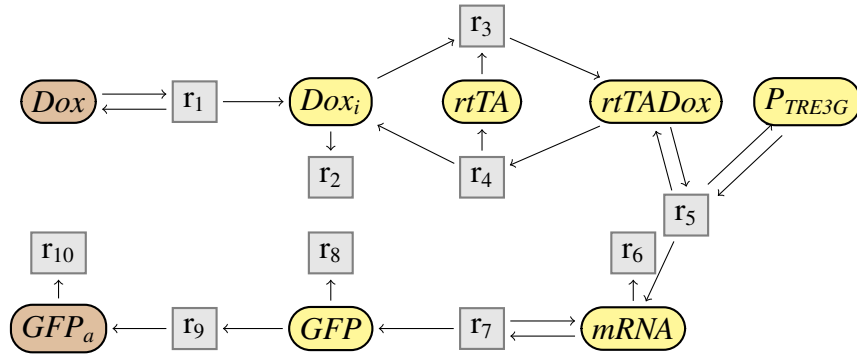


Fig 1.9: Réseau initial du système Tet-On.

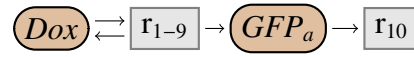


Fig 1.10: Réseau simplifié du système Tet-On.

tiellement simplifié illustré dans la Fig. 1.12, où la réaction  $r_7$  a été combinée respectivement avec  $r_8$  et  $r_9$ .

Un autre type de règles supprime une *réaction imitable*, c'est-à-dire lorsqu'on peut obtenir exactement la même transformation en utilisant un ensemble d'autres réactions du réseau. Une troisième catégorie se base sur l'existence de *symétries* dans le réseau pour pouvoir fusionner des espèces. Enfin, la simplification comprend d'autres règles plus diverses, permettant par exemple de supprimer des réactions qui ne changent pas l'état du réseau (comme par exemple la réaction  $r_{78}$  dans la Fig. 1.12), ou qui ne sont jamais applicables.

Nous considérons également un sous ensemble de ces règles de simplification, avec uniquement 6 règles. Nous prouvons d'une part que cet ensemble de règles est suffisant pour simplifier le réseau initial de Tet-On (Fig. 1.9), avec 10 réactions et 8 espèces, et obtenir le réseau simplifié présenté dans la Fig. 1.10, avec uniquement 2 réactions et 2 espèces. D'autre part, nous montrons que cet ensemble est *confluent*, et donc que quelque soit l'ordre d'application de ces règles de simplification, on obtient toujours le même réseau complètement simplifié.

Pour pouvoir résoudre le défi d'avoir une simplification correcte, nous proposons une *nouvelle sémantique contextuelle*, à partir de la sémantique opérationnelle. Elle est basée sur l'observation des attracteurs, dans tout contexte, et en négligeant certaines espèces considérées comme intermédiaires. Nous appellerons cette nouvelle sémantique la *sémantique des attracteurs*.

L'approche adoptée pour définir la sémantique des attracteurs est d'appliquer

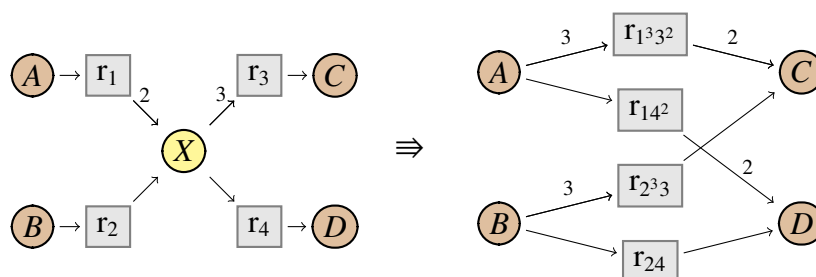


Fig 1.11: Exemple de simplification d'une espèce intermédiaire  $X$ . La réaction  $r_{1^3 3^2}$  est la combinaison de 3 fois la réaction  $r_1$  et de 2 fois la réaction  $r_3$ . Elle consomme trois molécules de  $A$ , produit 2 molécule de  $C$ , et ne modifie pas la quantité de  $X$ .

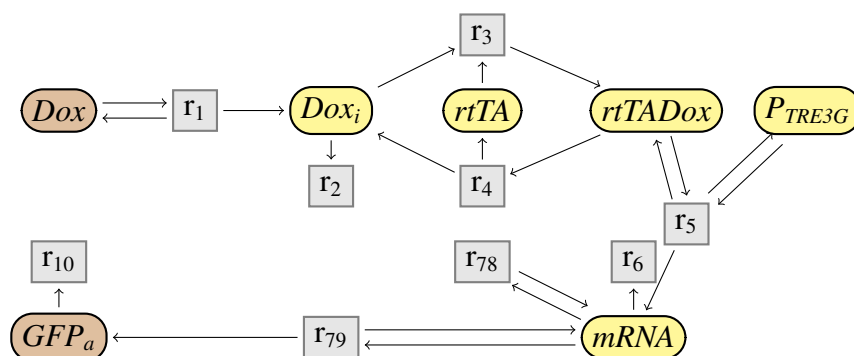


Fig 1.12: Réseau partiellement simplifié du système Tet-On, après avoir supprimé l'espèce intermédiaire  $GFP$ .

un cadre général pour définir des *sémantiques observationnelles* pour les langages de programmation, développé par [Schmidt-Schauss, Sabel, Niehren, and Schwinghammer \(2015\)](#), qui généralise la sémantique observationnelle définie avant, comme par exemple dans ([Pitts, 2002](#)). Ces sémantiques s'intéressent au comportement *entrée-sortie* d'un programme : placé dans un certain *contexte*, on observe si le programme peut (ou doit) *terminer*, ou au contraire peut diverger. En considérant un réseau de réactions comme un programme, il est alors possible d'adopter le même type de méthodes. Néanmoins, ce n'est pas direct, puisque la terminaison n'est pas un critère pertinent pour la biologie, où un réseau peut atteindre un ensemble d'états dans lequel il oscille, sans terminer.

Un *attracteur* (voir par exemple ([Naldi, Monteiro, & Chaouiya, 2012](#))) est un ensemble d'états, qui est fortement connexe et terminal. Cette notion est notamment bien connue dans les réseaux booléens ([Kauffman, 1969](#); [Thomas,](#)

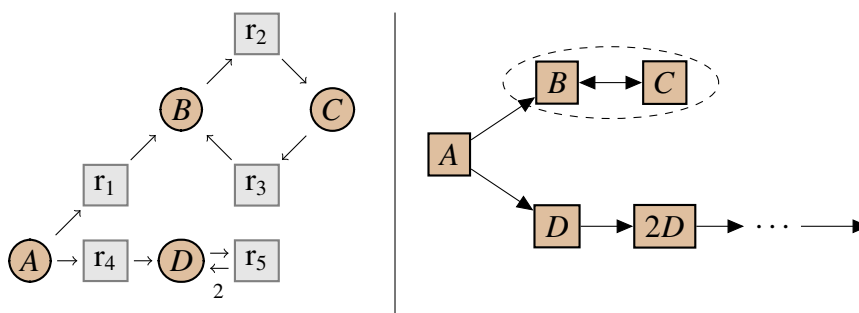


Fig 1.13: À gauche : graphe des réactions. À droite : graphe des état accessibles. Le réseau peut soit converger vers  $\{B, C\}$ , soit diverger.

1973) Il représente donc un comportement final et régulier d'un réseau puisqu'un réseau qui atteint un attracteur ne va jamais pouvoir le quitter, et va pouvoir atteindre n'importe quel état de l'attracteur. Cette notion peut être étendue aux graphes d'accessibilité des réseaux de réactions, en tenant compte du fait que les attracteurs peuvent alors être de taille infinie, et qu'un réseau peut aussi diverger, c'est-à-dire atteindre un état qui ne peut plus jamais converger vers un attracteur.

Par exemple, considérons le réseau de réactions présenté à gauche dans la Fig. 1.13. Le graphe des états accessibles avec ce réseau, en partant avec une unique molécule de  $A$ , est représenté à droite de la figure. Dans ce réseau, une molécule de  $A$  peut se transformer soit en une molécule de  $B$ , soit en une molécule de  $D$ . La molécule de  $B$  peut uniquement se transformer en  $C$ , et  $C$  peut uniquement se transformer en  $B$ . L'ensemble des états  $\{B, C\}$  forme alors un attracteur. La molécule de  $D$  peut alors se multiplier, créant deux molécules de  $D$ , puis trois, etc. On ne peut alors plus atteindre d'attracteur. Le réseau initial peut donc soit converger vers l'attracteur  $\{B, C\}$ , soit diverger en créant infiniment des molécules de  $D$ .

Nous considérons ici une notion d'attracteurs modulo une fonction d'observation, qui nous permet de ne pas prendre en compte les espèces intermédiaires lorsque de la comparaison de réseaux. Dans la *sémantique des attracteurs* que nous proposons, deux réseaux de réactions seront équivalents si, dans tout contexte, ils peuvent converger (ou diverger) vers les mêmes attracteurs, modulo cette observation. Nous prouvons alors que la simplification préserve cette sémantique, et donc que les deux réseaux de Tet-On sont bien équivalents.

Pour illustrer la pertinence biologique de notre méthode de simplification, nous l'appliquons sur différents modèles du système MAPK (Markevich et al., 2004), issus de la base de données de modèles [Biomodels](#). En les réduisant tous en un unique réseau, nous prouvons alors que les réseaux initiaux étaient tous équivalents pour la sémantique des attracteurs.

### 1.3.2 Simplification en déterministe

Pour la sémantique déterministe, la simplification de réseaux cinétiques doit surmonter la difficulté supplémentaire de calculer les taux cinétiques des réactions simplifiées, permettant de préserver exactement le comportement déterministe du réseau.

Nous proposons un ensemble de règles de *simplifications*, qui préservent de façon *exacte* la sémantique déterministe des réseaux de réactions. Ces règles sont structurelles et contextuelles. Elles permettent, notamment, de simplifier le réseau de réactions cinétiques de Tet-On, et d'obtenir les mêmes réactions que pour la simplification non déterministe, tout en calculant les expressions cinétiques des deux réactions simplifiées.

L'idée ici est d'*étendre les règles* de simplification précédentes, en ajoutant les cinétiques et en préservant la sémantique déterministe. Les règles calculent alors dans le même temps la structure des réactions simplifiées et leurs nouvelles cinétiques. Les règles restent donc *contextuelles*. Elles sont également purement *symboliques*, ne nécessitant pas de connaître la valeur des paramètres, ni de simuler numériquement le comportement du réseau. La simplification préserve le comportement déterministe de façon *exacte*, et non approchée.

Comme pour le cas non déterministe, des règles permettent alors de simplifier des *espèces intermédiaires*. Les règles calculent alors les nouvelles expressions cinétiques des réactions simplifiées. Il est alors nécessaire d'ajouter des deux restrictions supplémentaires par rapport au cas non déterministe. La première condition est que les espèces intermédiaires doivent être *à l'équilibre*, c'est-à-dire que leurs concentrations soient (exactement) constantes au cours du temps. Les autres espèces ne sont par contre pas obligatoirement à l'équilibre. De plus, il est nécessaire, sauf dans certains cas spécifiques, que les expressions cinétiques soient linéaires dans la concentration des espèces intermédiaires. Sous ces hypothèses, la règle d'élimination des espèces intermédiaires peut alors être formulée de manière plus générale que dans le cas non déterministe.

Comme pour le cas non déterministe, on peut également supprimer une *réaction imitable*, ou utiliser des *symétries entre espèces*. Certaines règles sont également spécifiques à la sémantique déterministe, par exemple en factorisant les réactants et les produits d'une réaction.

Considérons de nouveau le réseau Tet-On, cette fois avec des cinétiques. Il est représenté dans la Fig. 1.14. Toutes les cinétiques sont des lois d'action de masse. On peut alors appliquer la simplification jusqu'à obtenir le réseau simplifié de la Fig. 1.15. La structure du réseau simplifié est la même que pour la simplification en non déterministe. Le nombre de paramètres diminue de 11 initialement à 3

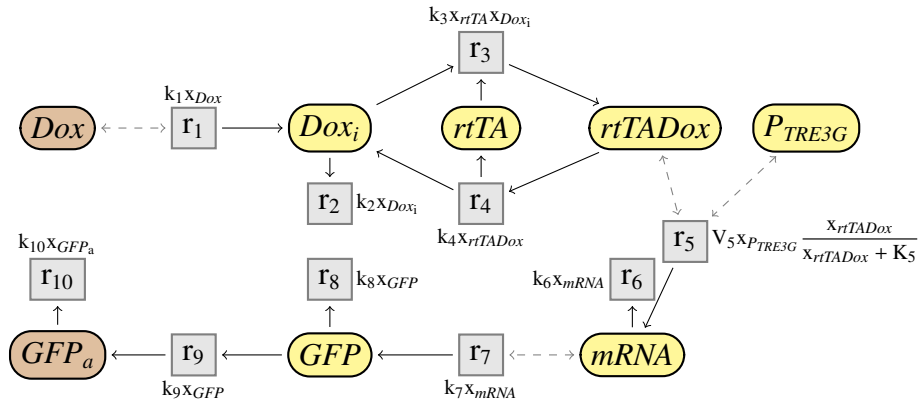


Fig 1.14: Réseau cinétique initial du système Tet-On.

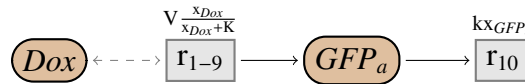


Fig 1.15: Réseau cinétique simplifié du système Tet-On.

dans le réseau simplifié.

### 1.3.3 Confluence

Enfin, nous étudions la confluence de la simplification des réseaux avec la sémantique déterministe, pour un sous-ensemble de règles, permettant en particulier d'éliminer des espèces intermédiaires. Plusieurs adaptations et restrictions sont alors nécessaires pour pouvoir obtenir un résultat de confluence.

En effet, la règle de simplification d'espèces intermédiaires, étant plus générale que dans le cas non déterministe, n'est pas confluente si considérée toute seule. Même en négligeant les cinétiques, on s'aperçoit que selon l'ordre d'élimination des espèces intermédiaires, des réactions supplémentaires peuvent parfois apparaître. Ces réactions, que nous appellerons plus tard *réactions dépendantes*, peuvent être supprimées à l'aide d'une nouvelle règle de simplification. Dans ce cas, et toujours sans cinétique, on obtient alors un résultat positif de confluence.

De plus, ce résultat est directement lié au calcul des *modes élémentaires* (Schuster & Hilgetag, 1994). Un mode est un *invariant*, un multi-ensemble de réactions, qui appliquées toutes ensemble, ne modifient pas la quantité des espèces intermédiaires. Un mode élémentaire vérifie en plus une notion de minimalité. Chaque

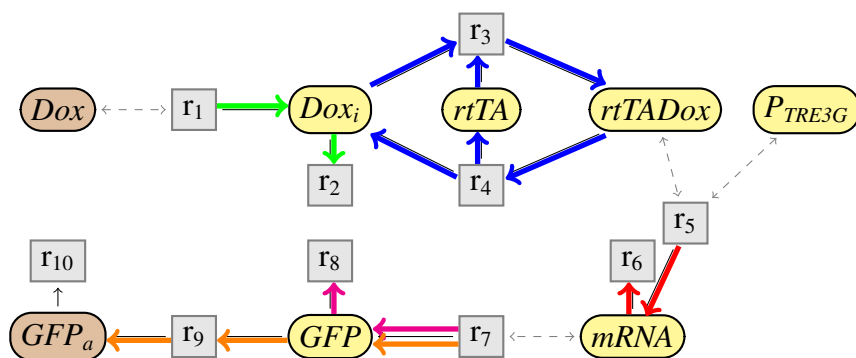


Fig 1.16: Modes élémentaires du système Tet-On (sans cinétique).

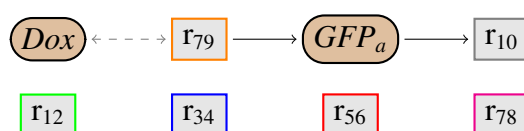


Fig 1.17: Réseau simplifié (sans cinétique) du système Tet-On.

réaction simplifiée produite par notre méthode est alors exactement la combinaison des réactions d'un mode élémentaire.

Les modes élémentaires du réseau Tet-On sont représentés dans la Fig. 1.16. Par exemple, le mode élémentaire  $\{r_3, r_4\}$  est représenté en bleu, et le mode élémentaire  $\{r_7, r_9\}$  est en orange. La simplification des espèces intermédiaires et des réactions dépendantes produit alors le réseau représenté dans la Fig. 1.17. Par exemple, la combinaison des réactions  $r_7$  et  $r_9$  forme la réaction simplifiée  $r_{79}$ , qui produit une molécule de  $GFP_a$ .

Lorsqu'on ajoute de nouveau les cinétiques, il est alors nécessaire d'ajouter des conditions afin de pouvoir éliminer l'ensemble des espèces intermédiaires. Notamment, les réactions consommant des espèces intermédiaires doivent avoir une *cinétique linéaire* dans leurs concentrations, et il ne peut pas y avoir plusieurs espèces intermédiaires en réactants (ou en produits) d'une réaction. Sous ces conditions, il est alors possible d'éliminer toutes les espèces intermédiaires, et les réseaux complètement simplifiés ont tous la *même structure*, c'est-à-dire les mêmes réactions lorsqu'on néglige les cinétiques.

Ce n'est néanmoins toujours pas suffisant pour avoir les mêmes cinétiques. On illustre en particulier ce problème sur un système biologique, pour lequel on peut trouver deux réseaux simplifiés avec la même structure mais des cinétiques



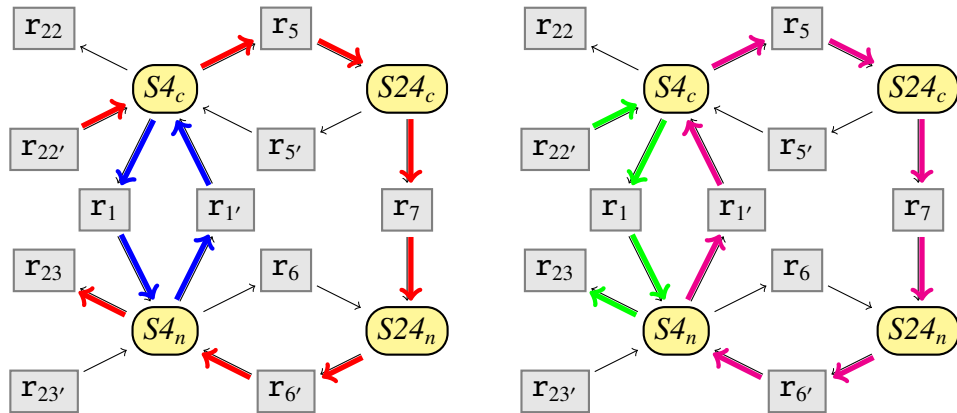


Fig 1.18: Réseau de signalisation pour l'espèce *Smad*, avec un mode décomposable en modes élémentaires de deux façons différentes.

différentes. On définit alors le *critère* suivant sur le réseau initial : un réseau est *uniquement décomposable* si, pour tout mode de ce réseau, il existe une unique décomposition de ce mode en des modes élémentaires.

Par exemple, considérons le réseau représenté dans la Fig. 1.18, représentant une sous partie du réseau de signalisation pour l'espèce *Smad* (Schmierer, Tournier, Bates, & Hill, 2008), où toutes les espèces sont intermédiaires. L'ensemble de réactions  $\{r_{22'}, r_5, r_7, r_{6'}, r_{23}, r_1, r_{1'}\}$  est alors un mode. Il peut se décomposer en modes élémentaires de deux façons différentes, soit avec les modes élémentaires  $\{r_{22'}, r_5, r_7, r_{6'}, r_{23}\}$  et  $\{r_1, r_{1'}\}$  représentés sur la figure de gauche, soit avec les modes  $\{r_{22'}, r_1, r_{23}\}$  et  $\{r_5, r_7, r_{6'}, r_{1'}\}$  sur la figure de droite. La simplification ne sera alors pas confluente pour les cinétiques sur ce réseau.

Lorsque le critère est vérifié, on prouve alors que la simplification est *complètement confluente*, à la fois pour la structure et les cinétiques. Par exemple, le réseau Tet-On vérifie le critère, et possède donc un unique réseau complètement simplifié.

## 1.4 Discussion

Il est alors naturel de se demander si on obtient des règles de simplification préservant à la fois la sémantique des attracteurs et la sémantique déterministe. Bien que le type de règles est le même pour les deux sémantiques, nous montrerons que la plupart du temps on ne peut pas complètement les unifier. En effet, les différences de traitement, par exemple des quantités de matières (nombre de molécules ou concentration), impliquent des conditions trop différentes dans les règles. En particulier, la règle d'élimination des espèces intermédiaires n'est pas unifiable.

Nous discuterons également brièvement de la possibilité d'étendre ces règles pour une sémantique stochastique. Nous montrerons que certaines règles peuvent trivialement être étendues. Néanmoins, la sémantique stochastique implique également des difficultés supplémentaires, notamment sur le type de contexte possible.

## 1.5 Plan de la thèse

**Chapitre 2 - Préliminaires** Ce premier chapitre introduit les différentes notions et définitions nécessaires à la compréhension de cette thèse. Elle définit d'abord les réseaux de réactions et la sémantique opérationnelle non déterministe, ainsi que la notion centrale d'observation. Ensuite, nous ajoutons les cinétiques aux réseaux, et décrivons la sémantique déterministe, à l'aide des systèmes d'équations différentielles. Enfin, nous donnons une définition formelle du type de simplification étudiée dans cette thèse, et de leurs propriétés.

**Chapitre 3 - Congruence des attracteurs et simplification non déterministe** Dans ce chapitre, nous définissons tout d'abord la sémantique des attracteurs pour les réseaux de réactions, représentant leurs comportements finaux réguliers. Nous étudions alors un premier ensemble de règles de simplification, confluent, permettant de simplifier le réseau biologique Tet-On (Huang et al., 2010). La simplification est alors enrichie avec d'autres règles, plus générales, mais non confluentes. Ces règles sont utilisées pour comparer et prouver l'équivalence de différents modèles du système MAPK (Markevich et al., 2004). Nous discutons alors de la sémantique des attracteurs et de notre méthode de simplification, en les comparant à la littérature.

Ce chapitre reprend et étend les résultats présentés dans (Madelaine, Lhoussaine, & Niehren, 2014).

**Chapitre 4 - Simplification préservant la sémantique déterministe** Dans ce chapitre, nous étudions la simplification avec une sémantique déterministe. On définit tout d'abord une notion de similarité, permettant de réécrire des expressions cinétiques en d'autres expressions équivalentes. Nous donnons ensuite un ensemble de règles de simplification pour la sémantique déterministe. Elles permettent d'obtenir un réseau simplifié ayant le même comportement que le réseau initial. Par exemple, des règles suppriment des espèces intermédiaires supposées à l'équilibre, ou utilisent des symétries du réseau pour fusionner certaines espèces. On étudie alors la simplification correspondante appliquée aux systèmes d'équations. Enfin, nous discutons de la simplification en la comparant à la littérature.

Ce chapitre reprend et étend les résultats présentés dans (Madelaine, Lhoussaine, & Niehren, 2015) et (Madelaine, Lhoussaine, Niehren, & Tonello, 2016).

**Chapitre 5 - Simplification confluyente des espèces intermédiaires linéaires** Ce chapitre s'intéresse plus particulièrement à la confluence de l'élimination des espèces intermédiaires en sémantique déterministe. Nous montrons que cette règle

de simplification, seule, n'est pas confluente, même en négligeant les cinétiques. En ajoutant deux nouvelles règles de simplification, nous prouvons néanmoins qu'on peut obtenir une simplification confluente (toujours sans les cinétiques). De plus, nous montrons également que cette simplification permet alors de calculer les modes élémentaires d'un réseau. Avec les cinétiques, et en ajoutant des conditions de linéarité sur les cinétiques, nous montrons que quelque soit l'ordre de la simplification, la structure du réseau complètement simplifié est toujours la même. En revanche, la répartition des taux cinétiques peut être différente, comme illustré sur un modèle biologique. Nous donnons alors un critère structurel sur le réseau initial suffisant pour garantir l'unicité de cette répartition, et donc la confluence de la simplification. Enfin, nous discutons de ce résultat en le comparant à la littérature. Ce chapitre reprend les résultats présentés dans (Madelaine, Tonello, Lhoussaine, & Niehren, 2016a) et dans (Madelaine, Tonello, Lhoussaine, & Niehren, 2016b).

**Chapitre 6 - Discussion** Enfin, dans ce dernier chapitre, nous discutons de la généralisation de nos règles de simplification à différentes sémantiques. Dans un premier temps, nous comparerons la sémantique non déterministe et la sémantique déterministe, et montrerons que leurs différences rendent parfois impossible une uniformisation des règles de simplification. Nous discutons alors brièvement de la généralisation de nos règles pour une sémantique stochastique, et des difficultés supplémentaires que cela implique, notamment pour la notion de contexte.



## CHAPITRE 2

---

### Préliminaires

---

Ce premier chapitre préliminaire introduit les concepts fondamentaux de la modélisation en biologie. On y définit formellement les différentes notions qui seront utilisées dans cette thèse, en les illustrant dans des exemples. On s'intéresse tout d'abord aux réseaux de réactions sans cinétique et à leur sémantique opérationnelle non déterministe. On étudie ensuite la sémantique déterministe des réseaux de réactions avec cinétiques et des systèmes d'équations différentielles ordinaires. Enfin, on présente les différentes propriétés que l'on souhaite avoir pour nos simplifications, c'est-à-dire des simplifications structurelles, correctes vis à vis d'une relation d'équivalence, et si possible confluentes.

**Notations** On note  $\mathbb{B} = \{true, false\}$  l'ensemble des booléens,  $\mathbb{N} = \{0, 1, \dots\}$  celui des entiers naturels,  $\mathbb{Z}$  celui des entiers relatifs,  $\mathbb{R}$  celui des nombres réels,  $\mathbb{R}_{\geq 0}$  celui des nombres réels positifs et  $\mathbb{R}_{> 0}$  celui des nombres réels strictement positifs.

## 2.1 Sémantique non déterministe des réseaux

Dans cette section, on définit formellement les réseaux de réactions, et leurs liens avec les réseaux de Pétri. On illustre ces définitions sur des exemples que nous réutiliserons tout au long de cette thèse. On présente la notion centrale d'observation, permettant de spécifier les espèces qui nous intéressent. On décrit ensuite la sémantique opérationnelle et non déterministe classique de ces réseaux.

### 2.1.1 Réseaux de réactions

**Réseaux de réactions** Une *espèce moléculaire*  $A, B, \dots \in \text{Espèces}$  est l'élément de base de nos modèles. Une espèce peut représenter biologiquement une protéine, une enzyme, un gène, de l'ARN messenger, etc. Une *solution chimique* est  $s : \text{Espèces} \rightarrow \mathbb{N}$  est un multi-ensemble de molécules. L'entier  $s(A)$  est appelé le *coefficient stœchiométrique* de l'espèce  $A$  dans la solution  $s$ . On utilise en général la notation  $s = a_1A_1 + \dots + a_nA_n$  pour désigner la solution  $s$  telle que pour tout  $1 \leq i \leq n$ ,  $s(A_i) = a_i$ , et  $s(A) = 0$  pour tout autre espèce  $A$ . L'ensemble des solutions constructibles avec les espèces  $\text{Espèces}$  est noté  $\text{Solutions}(\text{Espèces})$ . Comme l'ensemble des espèces est en général fixé, on utilise en général juste la notation  $\text{Solutions}$ .

Une *réaction (chimique)* est une transformation d'une solution chimique en une autre solution chimique. Biologiquement, une réaction peut représenter la liaison ou la désunion de deux molécules, le transport d'une molécule d'un compartiment vers un autre compartiment, la transcription d'ADN en ARN messenger, la traduction de l'ARN messenger en protéine, etc. Formellement, une réaction est une paire de solutions :

$$\text{Réactions} \quad \mathbf{r} ::= s_1 \rightarrow s_2$$

Les espèces présentes dans  $s_1$  sont appelées les *réactants* de  $\mathbf{r}$ , tandis que celles dans  $s_2$  sont appelées *produits* de  $\mathbf{r}$ . On appelle *vecteur stœchiométrique de réaction* le vecteur  $\text{stoic}_{\mathbf{r}} : \text{Espèces} \rightarrow \mathbb{Z}$  défini par  $\text{stoic}_{\mathbf{r}}(A) = s_2(A) - s_1(A)$ , qui décrit le nombre de molécules produites par  $\mathbf{r}$  (ou consommées, lorsque ce nombre est négatif). Une espèce  $A$  qui est présente à la fois dans les réactants et les produits, mais dont la quantité n'est pas modifiée par  $\mathbf{r}$ , c'est-à-dire telle que  $s_1(A) = s_2(A) > 0$  (et donc  $\text{stoic}_{\mathbf{r}}(A) = 0$ ), est appelée un *activateur* de  $\mathbf{r}$ . Notons qu'on ne considère ici que des réactions irréversibles, c'est-à-dire qui ne peuvent transformer que  $s_1$  en  $s_2$ , et ne peuvent pas faire la transformation inverse. Une réaction réversible sera donc modélisée par deux réactions, une dans chaque sens.

**Exemple 1.** Ainsi, la réaction  $\mathbf{r} = 2A + B \rightarrow C + B$  transforme deux molécules de l'espèce  $A$  en une molécule de l'espèce  $C$ , en présence d'une molécule de

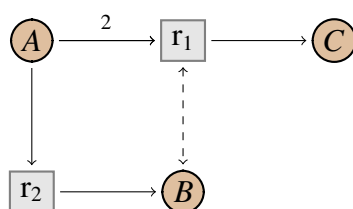


Fig 2.1: Représentation graphique des réactions  $r_1 = 2A + B \rightarrow C + B$  et  $r_2 = A \rightarrow B$ .

l'espèce  $B$ . Les réactants de  $r$  sont  $A$  et  $B$ , ses produits sont  $C$  et  $B$ . L'espèce  $B$  est un activateur dans  $r$ . La stœchiométrie de la réaction est égale à :

$$\text{stoic}_r(A) = -2 \quad \text{stoic}_r(B) = 0 \quad \text{stoic}_r(C) = 1.$$

Un ensemble de réactions  $R$  peut être représenté graphiquement, de façon similaire à des réseaux de Pétri. Il y a alors deux types de sommets : des sommets pour les espèces moléculaires, représentés par des cercles, et des sommets pour les réactions, représentés par des carrés. Un arc d'une espèce  $A$  vers une réaction  $r = s_1 \rightarrow s_2$  (resp. de  $r$  vers  $A$ ), labellisé par un entier  $i > 0$ , signifie que  $s_1(A) = i$  (resp.  $s_2(A) = i$ ). Lorsque  $A$  est un activateur pour  $r$ , c'est-à-dire  $s_1(A) = s_2(A) = i > 0$ , alors on utilise un arc  $A \leftrightarrow_i r$ . Lorsque  $i = 1$ , on omet en général le label.

**Exemple 2.** Les réactions  $R = \{r_1 = 2A + B \rightarrow C + B, r_2 = A \rightarrow B\}$  sont représentées graphiquement dans la Fig. 2.1.

Un *réseau de réactions*  $N = R \mid s_0$  est alors simplement une paire composée d'un ensemble de réactions  $R$  et d'une solution initiale  $s$ . La solution initiale représente l'état courant du réseau. On note  $R(N)$ , respectivement  $s_0(N)$ , les réactions et la solution initiale du réseau  $N$ . On note également  $\text{Espèces}(N)$  l'ensemble des espèces présentes soit dans une réaction soit dans la solution de  $N$ .

**Exemple 3.** Le réseau  $N = \{r_1 = 2A + B \rightarrow C + B, r_2 = A \rightarrow B\} \mid 3A$  possède deux réactions  $R(N) = \{r_1, r_2\}$ , et sa solution initiale  $s_0(N) = 3A$  contient trois molécules de l'espèce  $A$ . On a  $\text{Espèces}(N) = \{A, B, C\}$ .

**Réseau de Pétri** En biologie des systèmes, les réseaux de Pétri (Petri, 1966; Murata, 1989; Silva, 2013) sont souvent utilisés pour modéliser des systèmes biologiques (Reddy et al., 1993; Chaouiya, 2007; Heiner et al., 2008). Il y a en fait un lien direct entre les réseaux de réactions et les réseaux de Pétri.



**Définition 1.** Un *réseau de Pétri (marqué)* est un 4-uplet  $(P, T, A, m_0)$ , où :

- $P$  est un ensemble (fini) de *places*,
- $T$  est un ensemble (fini) de *transitions*,
- $A : (P \times T \cup T \times P) \rightarrow \mathbb{N}$  est une fonction associant une *multiplicité* à chaque *arc dirigé* entre les places et les transitions,
- $m_0 : P \rightarrow \mathbb{N}$  est une fonction associant à chaque place un nombre initial de *jetons*.

On a alors directement une équivalence entre un réseau  $N = R \mid s_0$  et le réseau de Pétri  $(P, T, A, m_0)$  défini par :

- l'ensemble des places sont les espèces :  $P = \text{Espèces}$ ,
- l'ensemble des transitions sont les réactions :  $T = R$ ,
- il existe un arc de  $A$  vers  $r$  (resp. de  $r$  vers  $A$ ) de multiplicité  $i$  si et seulement si  $i$  est la stœchiométrie de  $A$  dans les réactants (resp. produits) de  $r$ ,
- le marquage initial  $m_0$  représente la solution initial  $s_0$ .

On peut alors utiliser les résultats des réseaux de Pétri pour les réseaux de réactions. Par exemple, on peut noter que la plupart des équivalences usuelles (bisimulation, trace équivalence, etc.) ne sont pas décidables dans les réseaux de Pétri, et donc dans les réseaux de réactions (voir par exemple (Esparza, 1998) pour un résumé des résultats de complexité et de décidabilité dans les réseaux de Pétri). En revanche, le problème d'accessibilité est lui décidable (Mayr, 1984; Kosaraju, 1982), mais EXPSPACE-dur (Lipton, 1976).

**Contexte et composition** Dans cette thèse, on souhaite adopter une vision *modulaire* des réseaux. Ainsi, chaque réseau de réactions peut être vu comme un module, une sous-partie d'un réseau plus grand. On peut alors remplacer un module par un autre module, équivalent, sans changer le comportement global du réseau.

Par exemple, on peut étudier un réseau de réactions représentant, en présence d'un facteur de transcription, l'activation d'un gène et la production d'une protéine. Le facteur de transcription ou la protéine produite peuvent interagir avec d'autres éléments, qui ne nous intéressent pas pour ce modèle, et qui peuvent être inconnus. On considère alors que ces autres comportements font partie d'un contexte. Lorsque l'on cherchera ensuite à comparer des modèles, on considérera que deux modèles sont équivalents s'ils ont le même comportement dans tous les contextes possibles. Cette notion de contexte correspond à celle présente pour les langages

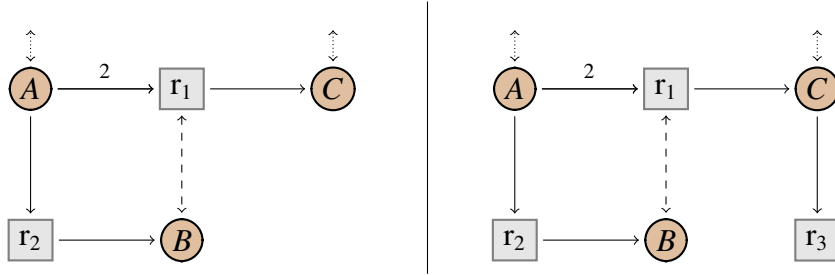


Fig 2.2: Représentation graphique du réseau  $N$  seul (à gauche) ou dans un contexte (à droite).

de programmation (Schmidt-Schauss et al., 2015). On voit ainsi un réseau comme un programme, qui reçoit des entrées (par exemple une production du facteur de transcription) et produit une sortie (la protéine). Notons que la sortie peut, selon le contexte, influencer sur l'entrée, et former ainsi une boucle de rétroaction.

On souhaite malgré tout que certaines des espèces soient spécifiques à notre réseau, et donc ne puisse pas interagir avec le contexte, comme par exemple le gène dans l'exemple ci-dessus. Formellement, ces espèces sont appelées *espèces internes*, et on note  $\mathcal{I} \subseteq \text{Espèces}$  l'ensemble de ces espèces. Cet ensemble sera en général fixé. Un *contexte* est alors un réseau  $N'$  qui n'utilise pas les espèces internes, c'est-à-dire  $\text{Espèces}(N') \cap \mathcal{I} = \emptyset$ . On définit alors la composition d'un réseau et d'un contexte par :

$$(R \mid s_0) \mid (R' \mid s'_0) = R \cup R' \mid s_0 + s'_0.$$

Lors de la représentation graphique des réseaux, on illustre les espèces pouvant interagir avec le contexte, c'est-à-dire les espèces non internes, par des flèches en pointillé.

**Exemple 4.** Reprenons le réseau de réactions précédent  $N = \{r_1, r_2\} \mid 3A$ . On suppose que les espèces  $A$  et  $C$  peuvent interagir avec le contexte, tandis que  $B$  est une espèce interne. Sa représentation graphique est à gauche de la Fig. 2.2.

Le réseau  $N' = \{r_3 = C \rightarrow \emptyset\} \mid A$ , qui ajoute une réaction  $r_3$  de dégradation pour  $C$  et une molécule de  $A$  dans la solution initiale, est alors un contexte valide, puisqu'il n'interagit pas avec l'espèce interne  $B$ . Le réseau  $N \mid N'$  est alors représenté à droite de la Fig. 2.2.

### 2.1.2 Sémantique opérationnelle

On va maintenant définir la *sémantique opérationnelle*, qui décrit comment un réseau évolue de façon non déterministe, de façon identique à la sémantique

usuelle des réseaux de Pétri. Si une réaction  $\mathbf{r} = s_1 \rightarrow s_2$  est présente dans un réseau, on peut alors l'appliquer et transformer la solution  $s_1$  en  $s_2$ . La relation de réduction est formellement définie par :

$$\frac{\mathbf{r} = s_1 \rightarrow s_2 \in \mathbf{R} \quad s_1 \subseteq s_0 \quad s'_0 = s_0 - s_1 + s_2}{(\mathbf{R} \mid s_0) \rightarrow_{\mathbf{r}} (\mathbf{R} \mid s'_0)}$$

On étend alors de façon naturelle la relation de réduction :

$$\rightarrow_{\mathbf{R}} = \bigcup_{\mathbf{r} \in \mathbf{R}} \rightarrow_{\mathbf{r}} \quad \rightarrow = \bigcup_{\mathbf{r} \in \text{Réactions}} \rightarrow_{\mathbf{r}}$$

**Exemple 5.** Reprenons le réseau  $N$  précédent, avec initialement trois molécules de  $A$ , et dont les réactions sont représentées graphiquement dans la Fig. 2.1. La réaction  $\mathbf{r}_1 = 2A + B \rightarrow C$  n'est pas applicable, puisqu'il n'y a pas de molécules de  $B$  dans la solution initiale. On peut par contre appliquer la réaction  $\mathbf{r}_2 = A \rightarrow B$ . La solution initiale  $s_0 = 3A$  est alors transformée en une solution  $s'_0 = 2A + B$ . On a donc  $N \rightarrow N'$ , avec  $N' = \mathbf{R} \mid s'_0$ .

À partir de  $N'$ , remarquons que les deux réactions sont cette fois applicables. Comme on utilise ici une sémantique non déterministe, on devra donc considérer les deux possibilités.

L'ensemble des réseaux accessibles à partir d'un réseau initial à l'aide de la sémantique opérationnelle peut être représenté à l'aide d'un graphe d'accessibilité.

**Définition 2** (Accessibilité et graphe d'accessibilité). Soit  $N$  un réseau. On définit l'ensemble des réseaux accessibles depuis  $N$  par :

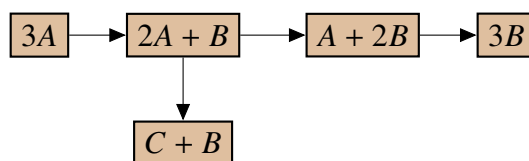
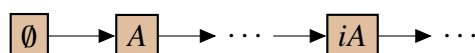
$$\{N \rightarrow^*\} = \{N' \mid N \rightarrow^* N'\}.$$

Le graphe d'accessibilité  $\mathcal{G}(N) = \langle \{N \rightarrow^*\}, \rightarrow, N \rangle$  est le graphe orienté et enraciné en  $N$  dont les sommets sont les réseaux accessibles, et dont les arcs correspondent à la relation de sémantique opérationnelle.

Le graphe des solutions accessibles  $\mathcal{G}_{sol}(N) = \langle V, E, v_0 \rangle$  est la projection de  $\mathcal{G}(N)$  sur les solutions, c'est-à-dire, si  $N = \mathbf{R} \mid s_0$  :

- $V = \{s \mid (\mathbf{R} \mid s) \in \{N \rightarrow^*\}\},$
- $E = \{(s, s') \mid (\mathbf{R} \mid s) \rightarrow (\mathbf{R} \mid s')\},$
- $v_0 = s_0.$

Pour tout graphe  $\mathcal{G} = \langle V, E \rangle$ , on notera  $\overline{\mathcal{G}}$  le graphe obtenu à partir de  $\mathcal{G}$  en supprimant les boucles, c'est-à-dire les arcs de forme  $(v, v)$ .

Fig 2.3: Graphe des solutions accessibles du réseau  $N$ .Fig 2.4: Graphe des solutions accessibles du réseau  $\{\emptyset \rightarrow A\} \mid \emptyset$ .

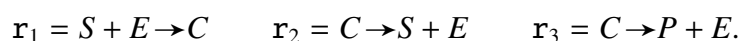
**Exemple 6.** Reprenons le réseau  $N = \{r_1, r_2\} \mid 3A$ , dont les réactions sont représentées dans la Fig. 2.1. Le graphe des solutions accessibles correspondant est représenté dans la Fig. 2.3.

Notons que l'ensemble des réseaux accessibles peut être de taille infinie, ainsi que le graphe d'accessibilité. Par exemple, le graphe des solutions accessibles du réseau  $\{\emptyset \rightarrow A\} \mid \emptyset$ , qui produit une infinité de molécules de  $A$ , est représentée dans la Fig. 2.4.

### 2.1.3 Exemples biologiques

Dans cette section, on considère trois exemples biologiques, que l'on retrouvera tout au long de cette thèse. Le premier exemple illustre la réaction enzymatique de Michaelis-Menten (sans cinétique) (Michaelis & Menten, 1913). On considère ensuite un modèle du système Tet-On (Gossen & Bujard, 1992; Gossen et al., 1995; Huang et al., 2010), et enfin un modèle de signalisation de MAPK (Markevich et al., 2004).

**Exemple 7** (Michaelis-Menten). La réaction enzymatique de Michaelis-Menten (Michaelis & Menten, 1913) représente la transformation d'un substrat  $S$  en un produit  $P$  en présence d'une enzyme  $E$ . Cette transformation peut être réalisée par un ensemble de trois réactions enzymatiques, ou par une unique réaction. Le réseau  $N_{MM} = \{r_1, r_2, r_3\} \mid S + E$  contient les 3 réactions suivantes, représentées graphiquement dans la Fig. 2.5 :



Dans la réaction  $r_1$ , le substrat  $S$  et l'enzyme  $E$  se lient pour former le complexe  $C$ . Par la réaction inverse  $r_2$ , le complexe  $C$  peut se décomposer en  $S$  et  $E$ . Il peut

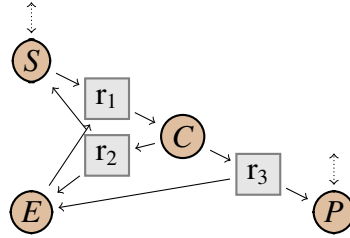


Fig 2.5: Graphe du réseau de Michaelis-Menten.

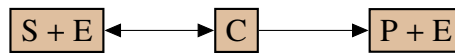


Fig 2.6: Graphe des solutions accessibles du réseau de Michaelis-Menten.

également, avec la réaction  $r_3$ , produire le produit  $P$ , tout en relâchant l'enzyme  $E$ .

Les espèces  $E$  et  $C$  sont ici des espèces internes, ne pouvant pas être modifiées par le contexte, contrairement aux espèces  $S$  et  $P$  :  $\mathcal{I} = \{E, C\}$ . La solution initiale du réseau comporte une molécule de  $S$  et une molécule de  $E$ .

Le graphe des solutions accessibles est représenté dans la Fig. 2.6.

On verra plus loin que cette transformation peut également se faire avec une unique réaction, transformant directement le substrat en produit.

**Exemple 8 (Tet-On).** On considère maintenant un réseau un peu plus gros du système Tet-On (Gossen & Bujard, 1992; Gossen et al., 1995; Huang et al., 2010). Il modélise une séquence de réactions qui, à partir de l'entrée dans la cellule d'un antibiotique, la doxycycline  $Dox$ , va déclencher l'expression d'un gène  $P_{TRE3G}$  et la production d'une molécule fluorescente  $GFP_a$ . Le modèle contient les espèces moléculaires suivantes :

$$Espèces = \{Dox, Dox_i, rtTA, rtTADox, P_{TRE3G}, mRNA, GFP, GFP_a\}.$$

Le réseau de réaction, représenté graphiquement sur la Fig. 2.7, est définie par :

$$N_{Tet-On} = R \mid Dox + 45000rtTA + 5P_{TRE3G}.$$

L'ensemble des espèces liées  $\mathcal{I} = Espèces \setminus \{Dox\}$  contient toutes les espèces, excepté la doxycycline  $Dox$ , qui est en dehors de la cellule et dont la quantité peut être contrôlée, par exemple en utilisant un dispositif micro-fluidique (Uhlendorf et al., 2012). La solution initiale du réseau contient une molécule de  $Dox$ , 45000

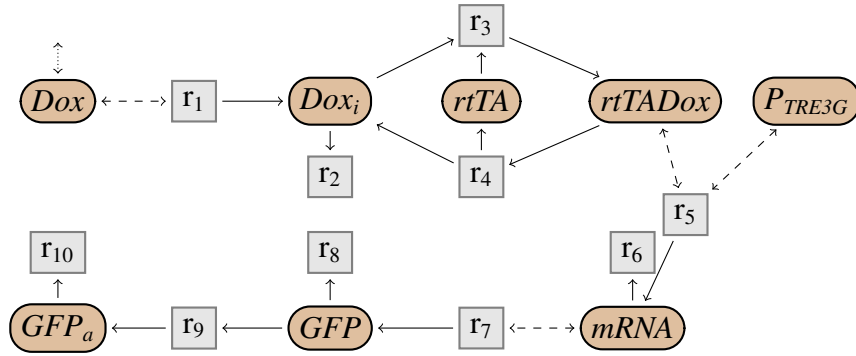


Fig 2.7: Graphe du système Tet-On.

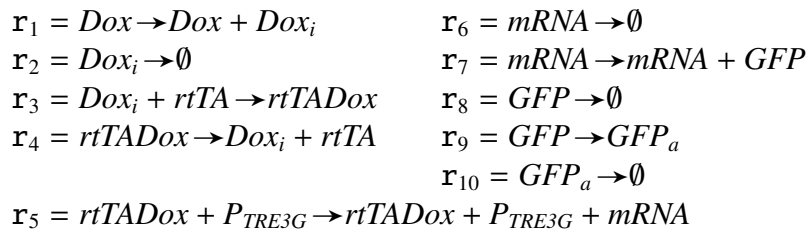


Fig 2.8: Réactions du réseau Tet-On.

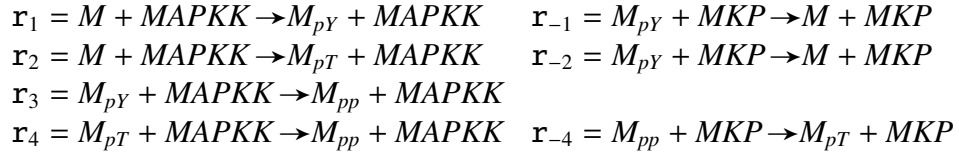


Fig 2.9: Réactions du système MAPK.

molécules de *rtTA*, et 5 copies du gène  $P_{TRE3G}$ . L'ensemble  $R$  de ces réactions est décrit formellement dans la Fig. 2.8.

Ce système modélise la production d'une protéine fluorescente verte ( $GFP_a$ ) dans une cellule, en présence d'un antibiotique, la doxycycline ( $Dox$ ), à l'extérieur de la cellule. On suppose ici que l'environnement peut contrôler la quantité de doxycycline à l'extérieur de la cellule (1 molécule dans la solution du réseau initial). Ainsi, l'entrée de doxycycline dans la cellule ne consomme pas  $Dox$ , qui est donc un activateur dans la réaction  $r_1$ . Une fois dans la cellule, la doxycycline  $Dox_i$  peut soit être dégradée par la réaction  $r_2$ , soit se lier au facteur de transcription artificiel *rtTA* par la réaction  $r_3$ . Le complexe  $rtTADox$  ainsi formé peut soit se dissocier ( $r_4$ ), soit activer la transcription du gène  $P_{TRE3G}$ , qui produit alors de l'ARN messager *mRNA* ( $r_5$ ). Celui-ci peut soit être dégradé par la réaction  $r_6$ , ou être traduit par la réaction  $r_7$  en protéine *GFP*. Enfin, la protéine peut devenir active et fluorescente ( $GFP_a$ ) avec la réaction  $r_9$ . Les protéines *GFP* et  $GFP_a$  peuvent également être dégradées par les réactions  $r_8$  et  $r_{10}$ .

**Exemple 9 (MAPK).** Enfin, on s'intéresse au système MAPK. Ce système décrit la double phosphorylation d'une molécule de  $M$ , créant la molécule  $M_{pp}$ . Dans (Markevich et al., 2004), plusieurs réseaux différents décrivant ce modèle sont proposés, selon le niveau de détails. On verra dans la suite comment on peut relier entre eux certains de ces modèles à l'aide de notre simplification. On présente ici le réseau *BIOMD0000000029* de la base de données de modèles biologiques [Biomodels](#). Il s'agit d'un réseau de taille intermédiaire proposé dans (Markevich et al., 2004) :

$$N_{MAPK} = R \mid M + a_1MAPKK + a_2MKP.$$

Le réseau contient une molécule de  $M$ , et des molécules de chaque activateur ( $MAPKK$  et  $MKP$ ). Les réactions sont décrites dans la Fig. 2.9 et représentées graphiquement dans la Fig. 2.10. On pose  $\mathcal{I} = \{M, M_{pp}, MKP, MAPKK\}$ .

Dans ce réseau, l'espèce moléculaire  $M$  possède deux sites,  $Y$  et  $T$ , pouvant être phosphorylés, sous l'action de l'activateur  $MAPKK$ . La réaction  $r_1$  phosphoryle le site  $Y$ , transformant  $M$  en  $M_{pY}$ , tandis que  $r_2$  transforme  $M$  en  $M_{pT}$ . Les réactions  $r_3$  et  $r_4$  phosphorylent le site restant, formant l'espèce  $M_{pp}$ ,

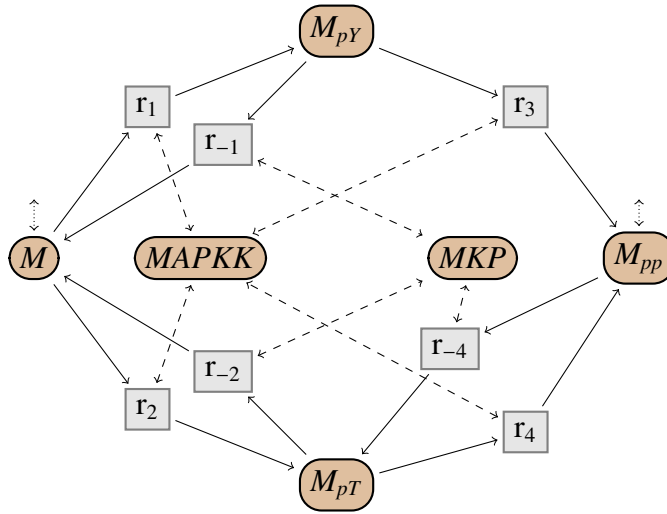


Fig 2.10: Graphe du système MAPK.

où les deux sites sont phosphorylés. Les sites peuvent être déphosphorylés sous l'action de l'activateur *MKP*. Notons que le site *T* ne peut être déphosphorylé que si le site *Y* n'est pas lui-même phosphorylé.

### 2.1.4 Observation

On va maintenant définir la notion centrale d'*observation*. Lors de la modélisation d'un système biologique, on est en général intéressé par l'évolution de certaines espèces moléculaires en particulier. Par exemple, considérons un modèle de la transcription puis traduction d'un gène. En général, on s'intéresse à la quantité de la protéine créée, plus qu'à celle de l'ARN messager, qui n'est une espèce intermédiaire. On ne veut donc observer que la protéine. On peut aussi ne connaître les données expérimentales que de certaines espèces. Enfin, on veut parfois ne pas faire la distinction entre plusieurs espèces, par exemple une protéine dans deux états différents (phosphorylée et non-phosphorylée).

Une *fonction d'observation*  $\Omega : \text{Espèces} \rightarrow \text{Observables}$  associe à chaque espèce moléculaire  $A$  une observation  $\Omega(A) \in \text{Observables}$ . L'ensemble *Observables* des observations peut être fini ou infini. On peut alors étendre cette fonction aux solutions, comme un morphisme du monoïde  $(\text{Solutions}, +, \emptyset)$  vers un monoïde commutatif d'observables  $(\text{Observables}, \oplus, \emptyset)$ , avec  $\emptyset$  l'observation vide :

$$\Omega(A) = \Omega(A) \quad \Omega(\emptyset) = \emptyset \quad \Omega(s + s') = \Omega(s) \oplus \Omega(s').$$

L'observation d'un réseau est alors l'observation de sa solution initiale :  $\Omega(\mathbb{R} | s) = \Omega(s)$ . En général, et sauf si précisé autrement, on fixera un ensemble fini



d'observations atomiques  $o \in ObsAtoms$ , et on utilisera comme monoïde celui des solutions construites à de ces observations atomiques. L'ensemble des observations est donc défini par  $Observables = Solutions(ObsAtoms)$ , la somme d'observations par la somme de solutions  $\oplus = +$ , et l'observation vide correspond à la solution vide  $\mathbb{O} = \emptyset$ . On a donc  $\Omega(s + s') = \Omega(s) + \Omega(s')$ .

**Exemple 10** (Observation de Michaelis-Menten). Dans la transformation enzymatique de Michaelis-Menten de l'Ex. 7, on s'intéresse principalement à la production de la protéine  $P$  en fonction de l'apport en substrat  $S$ . On va donc observer ces deux espèces. En revanche, la quantité d'enzymes,  $E$ , ne nous intéresse pas, donc cette espèce sera non observable. Le complexe  $C$  correspond à une molécule de  $E$  liée à une molécule de  $S$ . On va donc considérer que son observation est la somme de ses deux observations. On définit donc la fonction d'observation suivante sur les espèces, en prenant  $ObsAtoms = \{S, P\}$  :

$$\begin{aligned}\Omega_{MM}(S) &= S & \Omega_{MM}(P) &= P \\ \Omega_{MM}(E) &= \mathbb{O} & \Omega_{MM}(C) &= \Omega_{MM}(S + E) = S.\end{aligned}$$

On peut alors étendre cette observation aux solutions de façon classique, en considérant l'observation comme un morphisme de  $(Solutions, +, \emptyset)$  dans le monoïde  $(Solutions(\{S, P\}), +, \emptyset)$ , avec :

$$\Omega_{MM}(sS + eE + cC + pP) = (s + c)S + pP.$$

**Exemple 11** (Observation de Tet-On). Dans le système Tet-On décrit dans l'Ex. 8, on considère que seule la protéine  $GFP_a$ , qui est fluorescente, peut être observée. La fonction d'observation sur les espèces est alors définie par :

$$\Omega_{Tet-On}(A) = \begin{cases} GFP_a & \text{si } A = GFP_a, \\ \mathbb{O} & \text{sinon.} \end{cases}$$

On peut étendre cette fonction sur les solutions de différentes façons. On peut vouloir considérer la solution restreinte aux espèces observables (ici  $GFP_a$ ), et donc étendre la fonction d'observation de façon classique, avec :

$$\Omega_{Tet-On}(nGFP_a + s) = nGFP_a \quad \text{où } GFP_a \notin s.$$

On pourrait aussi être intéressé uniquement par la présence ou l'absence de  $GFP_a$ . On utiliserait alors le monoïde booléen  $(\{0, 1\}, max, 0)$  et le morphisme suivant :

$$\Omega'_{Tet-On}(nGFP_a + s) = min(n, 1) \quad \text{où } GFP_a \notin s.$$

**Exemple 12** (Observation de MAPK). Dans le système MAPK décrit dans l'Ex. 9, on veut connaître, pour chaque molécule, si elle possède 0, 1 ou 2 sites phosphorylés, sans savoir précisément quel site est phosphorylé. Ainsi,  $M_{pY}$  et  $M_{pT}$  ne seront pas distinguables. On a alors la fonction d'observation suivante sur les espèces :

$$\Omega_{MAPK}(A) = \begin{cases} A & \text{si } A \in \{M, M_{pp}, MAPKK, MKP\}, \\ M_p & \text{si } A \in \{M_{pY}, M_{pT}\}. \end{cases}$$

Comme discuté dans l'introduction, les notions d'équivalences usuelles pour la sémantique déterministe, comme par exemple la bisimulation, ne conviennent pas bien au cadre de la modélisation en biologie. Il est alors nécessaire de définir une nouvelle sémantique, basée sur les attracteurs et le comportement final des réseaux. Cette équivalence n'est pas présentée ici, elle sera décrite en détail dans le Chapitre 3.

## 2.2 Sémantique déterministe

Dans la section précédente, nous avons présenté les réseaux de réactions et leur sémantique non déterministe. Dans cette section, on étudie cette fois la sémantique déterministe. Chaque réaction du réseau possède une expression cinétique, représentant sa vitesse. On étudie alors, pour chaque espèce, l'évolution au cours du temps de sa concentration, à l'aide d'équations différentielles ordinaires. La concentration d'une espèce représente une quantité de molécules dans un volume fixé (donc l'espace d'états est ici continu). De même, la dynamique est désormais vue de façon continue au cours du temps, alors que l'on étudiait précédemment une sémantique étape par étape, donc à temps discret.

### 2.2.1 Réseaux de réactions cinétiques

On étend les réseaux de réactions en ajoutant à chaque réaction une expression cinétique, décrivant sa vitesse. On ajoute également des contraintes aux réseaux de réactions, permettant par exemple d'exprimer des conditions d'équilibre.

**Expressions** On note  $Vars$  un ensemble dénombrable de *variables*, que nous interpréterons plus tard comme des fonctions du temps vers les réels positifs. On suppose que pour chaque espèce moléculaire  $A \in Espèces$ , il existe une variable  $x_A \in Vars$  représentant sa concentration. On a également besoin de la valeur de sa concentration initialement, au temps  $t = 0$ . On suppose donc qu'il existe une variable (constante)  $x_A^0 \in Vars$ . On note  $Param$  un ensemble dénombrable de *paramètres*, qui seront interprétés comme des réels strictement positifs. Une *expression*  $e \in Expr$  est une composition de variables, de paramètres et de nombres réels, en utilisant les opérateurs mathématiques de base. La syntaxe des expressions est définie formellement dans la Fig. 2.11, avec  $x \in Vars$ ,  $k \in Param$  et  $a \in \mathbb{R}$ . On note  $Vars(e)$  l'ensemble des variables de  $e$ , et  $Espèces(e) = \{A \mid x_A \in Vars(e)\}$  l'ensemble de ses espèces. Notons que l'ensemble des espèces ne prend pas en compte la présence des concentrations initiales dans une expressions. Ainsi,  $A \notin Espèces(x_A^0)$ . Par commodité, on utilise des parenthèses ( $e$ ) quand la priorité des opérateurs n'est pas claire, et on définit les notations usuelles présentées dans la Fig. 2.11. Par exemple,  $k_1 x_A^2 x_B / (k_2 + x_D)$  est une expression, avec les paramètres  $k_1$  et  $k_2$ , et les concentrations des espèces  $A$ ,  $B$  et  $D$ .

Les expressions sont évaluées en assignant à chaque paramètre un réel strictement positif, et à chaque variable une fonction du temps vers une concentration, c'est-à-dire un réel positif. Soit  $\beta : Param \rightarrow \mathbb{R}_{>0}$  l'interprétation des paramètres. On suppose dans la suite que  $\beta$  est fixé afin de simplifier les notations, mais les simplifications seront correctes pour toute interprétation  $\beta$ . On

Syntaxe

$$e \in Expr ::= x \mid k \mid a \mid -e \mid e + e' \mid ee' \mid 1/e$$

Notations

$$e/e' = e(1/e') \quad e - e' = e + (-e) \quad e^n = \underbrace{e \dots e}_{n \text{ fois}}$$

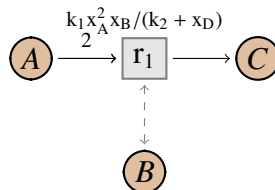
Interprétation

$$\begin{aligned} \llbracket x \rrbracket_\alpha &= \alpha(x) & \llbracket k \rrbracket_\alpha &= \lambda t \in \mathbb{R}_{\geq 0}. \beta(k) \\ \llbracket a \rrbracket_\alpha &= \lambda t \in \mathbb{R}_{\geq 0}. a & \llbracket -e \rrbracket_\alpha &= -\llbracket e \rrbracket_\alpha \\ \llbracket e + e' \rrbracket_\alpha &= \llbracket e \rrbracket_\alpha + \llbracket e' \rrbracket_\alpha & \llbracket ee' \rrbracket_\alpha &= \llbracket e \rrbracket_\alpha * \llbracket e' \rrbracket_\alpha \\ \llbracket 1/e \rrbracket_\alpha &= \begin{cases} 1/\llbracket e \rrbracket_\alpha & \text{si } \forall t \in \mathbb{R}_{\geq 0}. \llbracket e \rrbracket_\alpha(t) \neq 0 \\ undef & \text{sinon.} \end{cases} \end{aligned}$$

Fig 2.11: Expressions, avec  $x \in Vars$ ,  $k \in Param$ ,  $a \in \mathbb{R}$ .

note  $\alpha : Vars \rightarrow (\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0})$  l'interprétation des variables, vérifiant pour toute espèce  $A$  que sa concentration initiale  $x_A^0$  soit constante au cours du temps et égale à la concentration  $x_A$  au temps  $t = 0$ , c'est-à-dire  $\forall t. \alpha(x_A^0)(t) = \alpha(x_A)(0)$ . L'interprétation  $\llbracket e \rrbracket_\alpha \in (\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}) \cup \{undef\}$  d'une expression  $e \in Expr$  est alors décrite dans la Fig. 2.11. L'interprétation d'une expression  $e$  peut parfois être indéfinie, notée  $\llbracket e \rrbracket_\alpha = undef$ . C'est par exemple le cas pour l'expression  $1/e$  s'il existe un temps  $t$  tel que  $\llbracket e \rrbracket_\alpha(t) = 0$ .

**Réactions cinétiques et contraintes** Une *expression cinétique* est une expression  $e \in Expr$  toujours positive, c'est-à-dire telle que pour toute interprétation  $\alpha : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ , pour tout temps  $t$ , on a  $\llbracket e \rrbracket_\alpha(t) \geq 0$ . Une *réaction cinétique*  $r = s_1 \rightarrow s_2$ ;  $e$  est une paire composée d'une réaction  $s_1 \rightarrow s_2$  et d'une expression cinétique  $e$ . On notera  $kin(r) = e$  l'expression cinétique de la réaction  $r$ , qui représente la vitesse de la réaction au cours du temps. Lorsque qu'on représente graphiquement une réaction cinétique, on fait apparaître l'expression cinétique à coté du nœud représentant la réaction. Par exemple, la réaction  $r_1 = 2A + B \rightarrow C + B$ ;  $k_1 x_A^2 x_B / (k_2 + x_D)$  sera représentée par :



On appelle un *modificateur* de  $r$  une espèce  $A$  dont la quantité n'est pas modifiée par la réaction  $r$ , mais influe sur la vitesse de la réaction, c'est-à-dire telle que  $stoic_r(A) = 0$  et  $A \in Espèces(kin(r))$ . Par exemple,  $B$  et  $D$  sont deux modificateurs dans la réaction  $r_1$  précédente. L'augmentation de la concentration de  $B$  va ainsi accélérer la vitesse de la réaction, tandis que celle de  $D$  va la diminuer. Notons que la notion de modificateur est différente de celle d'activateur présentée précédemment. Par exemple,  $D$  ici n'est pas un activateur. Dans la sémantique non déterministe, une réaction n'est applicable que si ses activateurs (et ses autres réactants) sont présents dans la solution du réseau. En sémantique déterministe, elle est applicable au temps  $t$  si et seulement si l'interprétation de son expression cinétique au temps  $t$  est différente de zéro. La réaction précédente est ainsi applicable même si la concentration de  $x_D$  vaut zéro. En pratique, les activateurs sont souvent aussi des modificateurs, tandis que la réciproque est fautive. On verra un peu plus loin qu'en sémantique déterministe, on peut normaliser les réactions en supprimant les activateurs des réactants et des produits (sans modifier l'expression cinétique). Dans les représentations graphiques, les modificateurs d'une réaction seront en général indiqués avec une double flèche grisée et en pointillée. Ces flèches sont néanmoins optionnelles, puisque l'information est déjà présente dans l'expression cinétique de la réaction.

Une *contrainte*  $C \in Contraintes$ , dont la syntaxe est décrite dans la Fig. 2.12, ajoute des restrictions aux comportements possibles d'un réseau cinétique. Une contrainte est construite comme la conjonction  $\wedge$  de contraintes atomiques. La contrainte atomique  $e = e'$  impose que les expressions  $e$  et  $e'$  aient la même interprétation (et, en particulier, que leur interprétation soit bien définie). La contrainte  $cst(e)$  impose que l'interprétation de  $e$  soit constante au cours du temps. Les contraintes  $e \neq 0$  et  $e \geq 0$  imposent respectivement que  $e$  soit toujours différente de zéro, ou toujours supérieure à zéro. La contrainte  $\top$  est toujours vérifiée par définition. On définit également par commodité la contrainte  $e > 0 = (e \geq 0 \wedge e \neq 0)$ .

Une contrainte peut ainsi représenter le fait qu'une espèce  $A$  est à l'équilibre, c'est-à-dire que sa concentration est constante au cours du temps ( $cst(x_A)$ ). Elle peut aussi exprimer des lois de conservation, par exemple que la concentration totale de  $A$  et de  $B$  reste toujours la même ( $cst(x_A + x_B)$ ), des symétries entre des espèces ( $x_A = x_B$ ), ou encore des contraintes sur les conditions initiales ( $x_A^0 = k_1 x_B + k_2 x_C$ ).

L'interprétation  $\llbracket C \rrbracket_\alpha$  des contraintes, définie formellement dans la Fig. 2.12, attribue à chaque contrainte  $C$  une valeur dans  $\mathbb{B} \cup \{undef\}$ .

**Réseaux cinétiques** Un *réseau cinétique*  $M = R\&C$  est alors une paire avec  $R$  un ensemble de réactions cinétiques, et  $C$  une contrainte. On note  $R(M) = R$

Syntaxe

$$C \in \text{Contraintes} ::= cst(e) \mid e = e' \mid e \neq 0 \mid e \geq 0 \mid C \wedge C' \mid \top$$

Interprétation

$$\begin{aligned} \llbracket cst(e) \rrbracket_\alpha &= \exists c. \forall t. \llbracket e \rrbracket_\alpha(t) = c \\ \llbracket e = e' \rrbracket_\alpha &= \begin{cases} undef & \text{si } \llbracket e \rrbracket_\alpha = undef \text{ ou } \llbracket e' \rrbracket_\alpha = undef \\ \llbracket e \rrbracket_\alpha = \llbracket e' \rrbracket_\alpha & \text{sinon.} \end{cases} \\ \llbracket e \neq 0 \rrbracket_\alpha &= \forall t \in \mathbb{R}_{\geq 0}. \llbracket e \rrbracket_\alpha(t) \neq 0 & \llbracket C \wedge C' \rrbracket_\alpha = \llbracket C \rrbracket_\alpha \wedge \llbracket C' \rrbracket_\alpha \\ \llbracket e \geq 0 \rrbracket_\alpha &= \forall t \in \mathbb{R}_{\geq 0}. \llbracket e \rrbracket_\alpha(t) \geq 0 & \llbracket \top \rrbracket_\alpha = true \end{aligned}$$

Fig 2.12: Contraintes, avec  $e, e' \in Expr$ .

l'ensemble des réactions d'un réseau  $M$  et  $C(M) = C$  sa contrainte. On note également  $Expr(M)$  l'ensemble des expressions apparaissant soit en cinétique d'une réaction de  $R$ , soit dans la contrainte  $C$ . Enfin,  $Espèces(M)$  est l'ensemble des espèces  $A$  telles que soit  $A$  est dans une réaction, soit il existe une expression  $e \in Expr(M)$  avec  $A \in Espèces(e)$ .

De plus, on ne considérera que des réseaux cinétiques *normalisés*. Soit  $r = s_1 \rightarrow s_2$ ;  $e$  une réaction. On dit que  $r$  est *normalisée* si ses réactants et ses produits n'ont pas de molécules communes, c'est-à-dire  $s_1 \cap s_2 = \emptyset$ . En notant  $s = s_1 \cap s_2$ , on note  $\widehat{r} = s_1 - s \rightarrow s_2 - s$ ;  $e$  la normalisation de  $r$ . En effet, pour la sémantique déterministe, il n'est nécessaire que de connaître le vecteur de stœchiométrie  $stoic_r$ , et on a trivialement  $stoic_r(A) = stoic_{\widehat{r}}(A)$ . Par exemple, la réaction  $2A + B \rightarrow C + B$ ;  $k_1 x_A^2 x_B / (k_2 + x_D)$  est normalisée en  $2A \rightarrow C$ ;  $k_1 x_A^2 x_B / (k_2 + x_D)$ . Cette étape de normalisation n'est pas valable en sémantique non déterministe, puisqu'elle supprime les activateurs, et donc qu'une réaction normalisée n'est pas applicable de la même façon que la réaction originelle. En sémantique déterministe, on ne veut donc considérer que les réactions les plus simples pour ce vecteur, c'est-à-dire des réactions normalisées. Notons que dans la réaction ci-dessus, l'information sur le fait que l'espèce  $B$  doit nécessairement être présente pour pouvoir appliquer la réaction est déjà présente dans la cinétique : si  $x_B = 0$ , la vitesse de la réaction est directement égale à 0. Lors de la normalisation d'un réseau, on va également fusionner toutes les réactions ayant le même vecteur stœchiométrique, en sommant leurs cinétiques :

$$\widehat{R} = \{r \mid r \text{ normalisée}, kin(r) = \sum_{\substack{r' \in R, \\ stoic_{r'} = stoic_r}} kin(r')\}, \quad \widehat{R \& C} = \widehat{R} \& C.$$

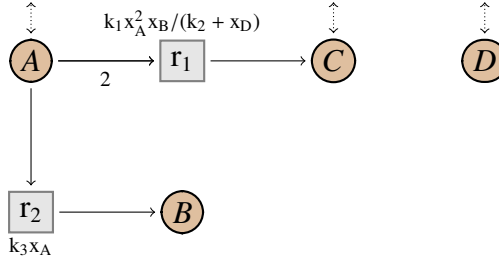


Fig 2.13: Représentation graphique du réseau cinétique  $M$ .

On note  $KinNets$  l'ensemble des réseaux cinétiques normalisés, et on ne considère dans la suite que des réseaux normalisés.

Comme pour les réseaux sans cinétiques, on considère qu'on a un ensemble  $\mathcal{I}$  fixé d'espèces internes. Alors, pour tout réseau  $M = R \& C$  et tout contexte  $M' = R' \& C'$  tel que  $Espèces(M') \cap \mathcal{I} = \emptyset$ , on définit leur composition par :

$$M \mid M' = R \cup R' \& C \wedge C'.$$

**Exemple 13.** On reprend l'exemple utilisé précédemment dans le cas non déterministe, en ajoutant des cinétiques et des contraintes :  $M = \{r_1, r_2\} \& C$ . Les réactions (normalisées)  $2A \rightarrow C$  ;  $k_1 x_A^2 x_B / (k_2 + x_D)$  et  $r_2 = A \rightarrow B$  ;  $k_3 x_A$  sont représentées graphiquement dans la Fig. 2.13. On suppose que seule l'espèce  $B$  est interne. On prend ici la contrainte  $C = k_2 + x_D \neq 0 \wedge cst(x_D)$ . Elle impose que l'on ne fait pas de division par zéro, donc que notre réseau soit bien défini, et que la concentration de  $D$  soit constante au cours du temps.

## 2.2.2 Systèmes d'équations

La sémantique déterministe d'un réseau cinétique va être exprimée à partir du système d'équations dérivé du réseau.

Une *équation différentielle ordinaire (ODE)* pour une variable  $x \in Vars$  est une équation  $\dot{x} = e$ , avec  $e \in Expr$  une expression. Elle est vérifiée par une interprétation  $\alpha$  si  $\llbracket e \rrbracket_\alpha$  est bien définie et égale à  $\llbracket \dot{x} \rrbracket_\alpha$ , où  $\dot{x}$  est la dérivée première de  $x$  par rapport au temps.

Un système d'équations  $E \in Systemes$  est la conjonction d'équations différentielles ordinaires et de contraintes, telle que pour chaque  $x$ , il y ait au plus une ODE pour  $x$  dans  $E$ . La syntaxe formelle des systèmes d'équations est définies dans la Fig. 2.15. On note respectivement  $Vars(E)$ ,  $Espèces(E)$  et  $Expr(E)$  les ensembles des variables, des espèces et des expressions du système  $E$ .

$$\begin{array}{ll}
\dot{x}_A = -2k_1x_A^2x_B/(k_2 + x_D) - k_3x_A & \wedge \quad \dot{x}_B = k_3x_A \\
\wedge \quad \dot{x}_C = k_1x_A^2x_B/(k_2 + x_D) & \wedge \quad \dot{x}_D = 0 \\
\wedge \quad k_2 + x_D \neq 0 & \wedge \quad cst(x_D).
\end{array}$$

Fig 2.14: Système d'équations pour le réseau  $M$ .

$$\begin{array}{l}
\text{Systèmes d'équations: } E ::= \dot{x} = e \mid C \mid E \wedge E' \\
\text{Interprétation: } \llbracket \dot{x} = e \rrbracket_\alpha = \begin{cases} \llbracket \dot{x} \rrbracket_\alpha = \llbracket e \rrbracket_\alpha & \text{si } \llbracket e \rrbracket_\alpha \neq \text{undef} \\ \text{undef} & \text{sinon} \end{cases} \\
\llbracket E \wedge E' \rrbracket_\alpha = \llbracket E \rrbracket_\alpha \wedge \llbracket E' \rrbracket_\alpha.
\end{array}$$

Fig 2.15: Syntaxe et interprétation d'un système d'équations.

**Exemple 14.** Le système d'équations  $E$  représenté dans la Fig. 2.14 contient quatre équations différentielles et deux contraintes.

Un système est satisfait si toutes ses *ODEs* et ses contraintes sont satisfaites, comme décrit dans la Fig. 2.15.

Comme pour le cas non déterministe, on considère une fonction d'observation  $\Omega : \text{Espèces} \rightarrow \text{Observables}$  sur les espèces. Comme précédemment, on prend un ensemble d'observations atomiques  $\text{ObsAtoms}$ , et on pose ici  $\text{Observables} = \text{ObsAtoms} \cup \{\emptyset\}$ . Ainsi, l'observation d'une espèce est soit une observation atomique, soit l'observation vide.

Pour toute  $\alpha : \text{Espèces} \rightarrow (\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0})$ , on définit l'interprétation observable  $\Omega(\alpha) : \text{ObsAtoms} \rightarrow (\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0})$  telle que pour tout  $o \in \text{ObsAtoms}$ , on ait :

$$\Omega(\alpha)(o) = \sum_{A \in \Omega^{-1}(o)} \alpha(x_A).$$

**Exemple 15.** Supposons qu'on ait un système avec cinq espèces  $\{A, A_1, A_2, B, C\}$ , deux observations atomiques  $o_A$  et  $o_B$ , et la fonction d'observation suivante :

$$\Omega(A) = \Omega(A_1) = \Omega(A_2) = o_A \quad \Omega(B) = o_B \quad \Omega(C) = \emptyset.$$

Alors on observe la somme des concentrations de  $A, A_1$  et  $A_2$ , ainsi que la concentration de  $x_B$ , mais pas  $x_C$ , ni la concentration de  $A, A_1$  ou  $A_2$  individuellement.

Pour toute fonction d'observation  $\Omega$ , on définit l'ensemble des *solutions observables d'un système*  $E$  comme l'ensemble des interprétations observables  $\Omega(\alpha)$  telles que  $\alpha$  satisfont le système  $E$  :

$$\text{sol}_\Omega(E) = \{\Omega(\alpha) \mid \llbracket E \rrbracket_\alpha = \text{true}\}.$$

Notons que l'on ne considère que des interprétations  $\alpha$  qui sont bien définies pour toutes les expressions  $e \in \text{Expr}(E)$ .



$$\begin{array}{l|l}
\dot{x}_{A_1} = k_1 - k_2 x_{A_1} & \dot{x}_{A_1} = 0 \\
\wedge \dot{x}_{A_2} = k_1 - k_2 x_{A_2} & \wedge \dot{x}_{A_2} = 0 \\
\wedge \dot{x}_A = 0 & \wedge \dot{x}_A = 2k_1 - k_2 x_A \\
\wedge \dot{x}_B = k_2(x_{A_1} + x_{A_2}) & \wedge \dot{x}_B = k_2 x_A \\
\wedge \dot{x}_C = k_3 x_B - x_C & \wedge \dot{x}_C = 0 \\
\wedge x_{A_1}^0 = 0 \wedge x_{A_2}^0 = 0 \wedge x_A^0 = 0. & \wedge x_{A_1}^0 = 0 \wedge x_{A_2}^0 = 0 \wedge x_A^0 = 0.
\end{array}$$

Fig 2.16: Ces deux systèmes sont équivalents, si on observe uniquement la somme des concentration de  $A$ ,  $A_1$  et  $A_2$ , et la concentration de  $B$ .

**Définition 3** (Équivalence de systèmes d'équations). Deux systèmes  $E$  and  $E'$  sont  $\Omega$ -équivalents, noté  $E \simeq_{\Omega} E'$ , s'ils ont les mêmes solutions observables :

$$E \simeq_{\Omega} E' \quad \text{si et seulement si} \quad \text{sol}_{\Omega}(E) = \text{sol}_{\Omega}(E').$$

**Exemple 16.** On reprend la fonction d'observation de l'exemple précédent. Alors les deux systèmes d'équations de la Fig. 2.16 sont équivalents.

En effet, remarquons d'abord que puisque qu'on observe pas du tout  $C$ , et que sa concentration n'intervient pas dans les *ODEs* des autres espèces, les systèmes seront équivalents quelque soit sa valeur.

Dans le réseau de gauche, les espèces  $A_1$  et  $A_2$  ont exactement le même comportement. On les fusionne en une unique espèce  $A$  dans le réseau de droite. Comme dans le réseau de gauche, on observe uniquement la somme des concentrations de  $A_1$  et  $A_2$ , et que le comportement de  $x_B$  ne dépend aussi que de cette somme, on obtient bien le même comportement observable.

### 2.2.3 Sémantique déterministe des réseaux cinétiques

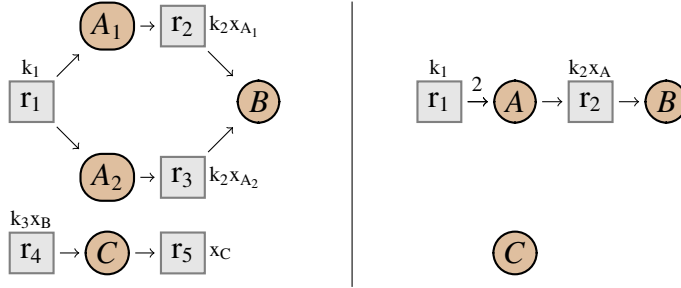
À partir de tout réseau cinétique, on peut alors dériver un système d'équations, et l'utiliser pour définir la sémantique déterministe du réseau.

Soit  $M = R\&C$  un réseau, on construit le système  $E(M)$  en conservant les contraintes et en calculant, pour chaque espèce moléculaire  $A$ , son équation différentielle ordinaire à l'aide des réactions. La construction est décrite formellement dans la Fig. 2.17. Remarquons qu'il y a une équation différentielle pour chaque espèce  $A \in \text{Espèces}$ , même lorsque  $A \notin \text{Espèces}(M)$ . Dans ce cas, son *ODE* est  $\dot{x}_A = 0$ .

**Exemple 17.** Le système d'équation dérivé du réseau cinétique  $M$  (Fig. 2.13) est celui de la Fig. 2.14.

$$E(R \mid C) = \bigwedge_{A \in \text{Espèces}} \dot{x}_A = \sum_{r \in R} \text{stoic}_r(A) \text{kin}(r) \wedge C$$

Fig 2.17: Dérivation du système d'équations d'un réseau cinétique.

Fig 2.18: Ces deux réseaux sont équivalents, si on observe uniquement la somme des concentration de  $A$ ,  $A_1$  et  $A_2$ , et la concentration de  $B$ .

**Définition 4** (Équivalence de réseaux cinétiques). Deux réseaux  $M$  et  $M'$  sont  $\Omega$ -équivalents, noté  $M \simeq_{\Omega} M'$ , si leurs systèmes d'équations sont  $\Omega$ -équivalents :

$$M \simeq_{\Omega} M' \quad \text{si et seulement si} \quad E(M) \simeq_{\Omega} E(M').$$

**Exemple 18.** On reprend la fonction d'observation précédente. Alors les deux réseaux de la Fig. 2.18, dont les systèmes sont ceux de la Fig. 2.16, sont équivalents.

Comme pour le cas non déterministe, on étend cette équivalence à une congruence contextuelle.

**Définition 5** (Congruence de réseaux cinétiques). Deux réseaux  $M$  et  $M'$  sont  $(\Omega, \mathcal{I})$ -congruents, noté  $M \sim_{\Omega, \mathcal{I}} M'$ , si pour tout contexte  $M''$  qui vérifie que  $\text{Espèces}(M'') \cap \mathcal{I} = \emptyset$ , on a  $M \mid M'' \simeq_{\Omega} M' \mid M''$ .

## 2.2.4 Exemples biologiques

On reconsidère ici les systèmes biologiques de Michaelis-Menten et de Tet-On présentés précédemment, cette fois avec des cinétiques.

**Exemple 19** (Michaelis-Menten). On définit le réseau cinétique pour le système de Michaelis-Menten (Michaelis & Menten, 1913) par :

$$M_{\text{MM}} = \{r_1, r_2, r_3\} \& \text{cst}(x_E) \wedge \text{cst}(x_C).$$

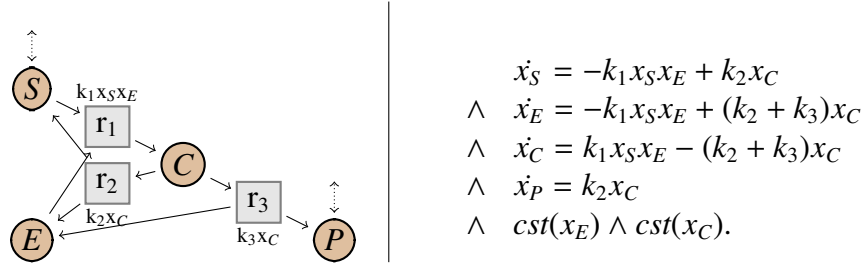


Fig 2.19: Graphe et système du réseau cinétique de Michaelis-Menten.

Les réactions sont représentées graphiquement à gauche de la Fig. 2.19, et le système d'équations correspondant à droite.

La contrainte impose ici que  $E$  et  $C$  soient à l'équilibre, c'est-à-dire que leurs concentrations soient constantes au cours du temps. Dans la littérature (voir (Palsson, 1987; Segel, 1988; Schnell & Mendoza, 1997; Sanft, Gillespie, & Petzold, 2011)), la simplification est en général une approximation, basée sur l'une des hypothèses suivantes : soit que la concentration de  $C$  est quasi-stable, soit que les vitesses des réactions  $r_1$  et  $r_2$  sont quasi à l'équilibre. Comme on souhaite ici une équivalence (et une simplification) exacte, la contrainte est plus forte : on demande à ce que la concentration de  $C$  soit exactement à l'équilibre.

Notons que la structure des réactions imposent que la concentration totale de  $E$  et de  $C$  soit constante. En effet, chaque réaction consommant (resp. produisant) une molécule de  $E$  produit également (resp. consomme) une molécule de  $C$ , et réciproquement. Comme de plus ces deux espèces sont internes, cette propriété reste vraie dans tout contexte. Donc la contrainte  $\text{cst}(x_E)$  découle naturellement de la contrainte  $\text{cst}(x_C)$ . On peut également le voir dans le système d'équation, puisqu'on a  $\dot{x}_E = -\dot{x}_C$ .

**Exemple 20 (Tet-On).** Considérons maintenant le réseau cinétique pour Tet-On (Gossen & Bujard, 1992; Gossen et al., 1995; Huang et al., 2010), qui étend le réseau de l'Ex. 8. Les réactions  $R$  du réseau initial  $M_{\text{Tet-On}} = R\&C$  sont représentées dans la Fig. 2.20. La contrainte est définie par :

$$C = \bigwedge_{A \in \text{Espèces} \setminus \text{GFP}_a} \text{cst}(x_A).$$

Donc on considère que toutes les espèces à part  $\text{GFP}_a$  sont à l'équilibre.

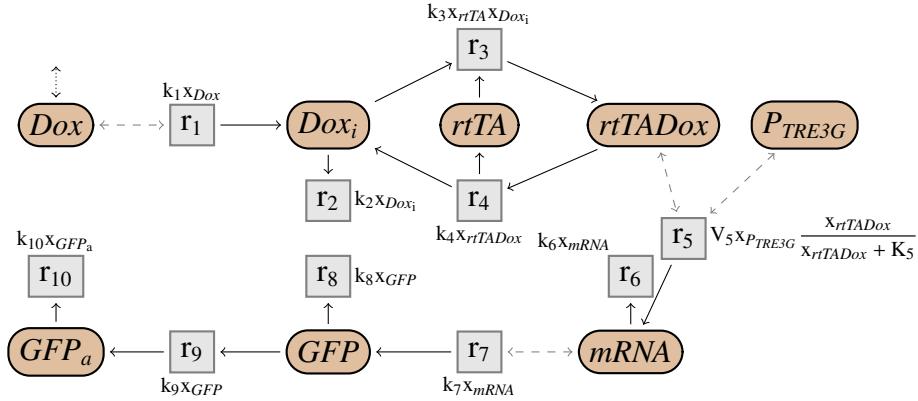


Fig 2.20: Graphe du réseau cinétique Tet-On.

## 2.3 Simplification

Cette thèse vise à simplifier des réseaux de réactions chimiques. Dans cette section, on définit formellement la forme de nos simplifications et leurs différentes propriétés.

### 2.3.1 Relation de simplification structurelle

Soit  $S$  un ensemble de systèmes, sur lesquels on veut appliquer des simplifications. On suppose que l'ensemble  $S$  est muni d'une relation d'ordre (partielle)  $<$ , ainsi que d'une relation d'équivalence structurelle  $\equiv$ . Par exemple,  $S$  peut être l'ensemble des réseaux de réactions chimiques. Dans ce cas, l'ordre peut être défini en fonction du nombre d'espèces et de réactions d'un réseau, et l'équivalence structurelle est l'égalité des réseaux. Ou bien  $S$  peut être l'ensemble des systèmes d'équations, et  $\equiv$  une relation permettant de réécrire des expressions, par exemple en les factorisant.

De plus, la relation d'ordre doit vérifier les deux propriétés suivantes :

- *préservation par l'équivalence structurelle* : si  $s_1 < s_2$ ,  $s_1 \equiv s'_1$  et  $s_2 \equiv s'_2$ , alors  $s'_1 < s'_2$ ,
- *bien fondée* : il n'existe pas de suite infinie de systèmes  $s_1, s_2, \dots, s_i, \dots$  telle que pour  $i$ ,  $s_{i+1} < s_i$ .

**Définition 6.** Une relation de simplification sur  $S$  est une relation  $\Rightarrow \in S \times S$  telle que pour tout  $(s, s') \in \Rightarrow$ ,  $s' < s$ . En d'autres termes, une relation de simplification va transformer un système  $s$  en un système  $s'$  plus petit. On utilisera en général la notation infixe pour les relations de simplification, c'est-à-dire  $s \Rightarrow s'$ .

$$\frac{\text{Conditions d'application}}{s \Rightarrow s'} \quad (\text{NOM DE LA RÈGLE})$$

Fig 2.21: Schéma d'une règle de simplification.

Dans cette thèse, nos simplifications sont définies par un ensemble de règles *structurelles*, qui suivent le schéma général décrit dans la Fig. 2.21, où un système  $s$  est simplifié en un système  $s'$ , sous certaines conditions. Ces règles s'appliquent en général directement sur la structure des systèmes, c'est-à-dire qu'elles ne nécessitent pas de calculer ou de simuler la dynamique.

### 2.3.2 Simplification correcte

En outre, on souhaite que nos simplifications préservent certaines propriétés sémantiques de nos réseaux. Ainsi, on définit une relation d'équivalence  $\sim$  sur les réseaux (par exemple, pour la sémantique déterministe, les deux réseaux doivent avoir les mêmes solutions observables), et on souhaite que le réseau simplifié soit équivalent au réseau initial. On dit alors que la simplification est *correcte* vis à vis de l'équivalence.

**Définition 7.** Une relation de simplification  $\Rightarrow \in S \times S$  est *correcte* pour une relation d'équivalence  $\sim \in S \times S$  si  $s \Rightarrow s'$  implique  $s \sim s'$ .

On peut également définir une notion de *simulation* entre deux relations de simplification, décrivant le fait qu'une simplification peut en imiter une autre. On peut alors utiliser les propriétés d'une simplification pour prouver des résultats sur l'autre. Considérons deux ensembles de systèmes  $S_1$  et  $S_2$  munis respectivement des relations d'équivalence  $\sim_1$  et  $\sim_2$ . Soit  $\Rightarrow_1 \in S_1 \times S_1$  et  $\Rightarrow_2 \in S_2 \times S_2$  deux relations de simplification. Une transformation  $T : S_1 \rightarrow S_2$  est une *simulation* de  $(S_1, \sim_1, \Rightarrow_1)$  dans  $(S_2, \sim_2, \Rightarrow_2)$  si, pour tout  $s, s' \in S_1$  :

- les équivalences correspondent :  $s \sim_1 s'$  si et seulement si  $T(s) \sim_2 T(s')$ ,
- les simplifications dans  $S_1$  sont imitables dans  $S_2$  :  $s \Rightarrow_1 s'$  implique  $T(s) \Rightarrow_2 T(s')$ .

Ces conditions sont représentées dans la Fig. 2.22. On verra par exemple que la dérivation d'un système d'équations à partir d'un réseau de réactions cinétiques est une simulation avec leurs relations de simplification respectives.

On a alors la propriété suivante :

**Proposition 1.** S'il existe une simulation  $T$  de  $(S_1, \sim_1, \Rightarrow_1)$  dans  $(S_2, \sim_2, \Rightarrow_2)$ , et que la simplification  $\Rightarrow_2$  est correcte pour  $\sim_2$ , alors la simplification  $\Rightarrow_1$  est correcte pour  $\sim_1$ .

$$\begin{array}{ccc|ccc}
s & \sim_1 & s' & s & \Rightarrow_1 & s' \\
\downarrow T & & \downarrow T & \downarrow T & & \downarrow T \\
T(s) & \sim_2 & T(s') & T(s) & \Rightarrow_2 & T(s')
\end{array}$$

Fig 2.22: Diagrammes de la simulation  $f$ .

*Proof.* Soit  $s, s' \in S_1$  tels que  $s \Rightarrow_1 s'$ , et soit  $T$  la simulation. On a donc  $T(s) \Rightarrow_2 T(s')$ , et comme  $\Rightarrow_2$  est correcte, cela implique  $T(s) \sim_2 T(s')$ , et donc  $s \sim_1 s'$ .  $\square$

### 2.3.3 Confluence

Enfin, on s'intéresse également à la notion de *confluence* d'une simplification (Huet, 1980). Avoir un système de simplification confluent va impliquer que quelque soit l'ordre dans lequel les règles de simplification sont appliquées, le système obtenu lorsqu'on ne peut plus le simplifier davantage sera toujours le même (modulo éventuellement l'équivalence structurelle). On va donc pouvoir appliquer par exemple les règles de simplification dans un ordre optimal d'un point de vue de la complexité, en garantissant l'obtention du réseau minimal à la fin de la simplification. On s'intéresse toujours ici à des notions de confluence modulo l'équivalence structurelle  $\equiv$ .

Si  $\Rightarrow \in S \times S$  est la relation de simplification, on note  $\Leftarrow$  la relation symétrique définie par  $\Leftarrow = \{(s, s') \mid (s', s) \in \Rightarrow\}$ . La clôture symétrique est définie par  $\Leftrightarrow = \Rightarrow \cup \Leftarrow$ . On pose  $\Rightarrow^0 = \equiv$ , et pour tout  $k > 0$ ,  $\Rightarrow^k = \Rightarrow^{k-1} \circ \Rightarrow$ . La clôture transitive est définie par  $\Rightarrow^+ = \bigcup_{k \geq 1} \Rightarrow^k$ , et la clôture réflexive transitive par  $\Rightarrow^* = \bigcup_{k \geq 0} \Rightarrow^k$ . Enfin,  $\Rightarrow^\epsilon = \Rightarrow^+ \cup \Rightarrow^0$ .

On dit alors qu'une simplification  $\Rightarrow$  est (*globalement*) *confluente* pour  $(S, \equiv)$  si  $\Leftarrow^* \circ \Rightarrow^* \subseteq \Rightarrow^* \circ \Leftarrow^*$ , comme illustré dans la Fig. 2.23. Ainsi, si un système  $s$  peut être simplifié en un système  $s_1$  ou  $s_2$  (en zéro, une ou plusieurs étapes de simplification), alors les systèmes  $s_1$  et  $s_2$  peuvent être simplifiés en un même système  $s'$ .

Afin de prouver la confluence, on utilisera en général la *confluence locale*. Une simplification est localement confluente (pour  $(S, \equiv)$ ) lorsque  $\Leftarrow \circ \Rightarrow \subseteq \Rightarrow^* \circ \Leftarrow^*$ , c'est-à-dire si on peut simplifier un système en une seule étape en deux systèmes différents, alors on peut ensuite les simplifier en un même système (en zéro, une ou plusieurs étapes) (voir Fig. 2.23). Le lemme de Newman indique que toute relation qui termine et qui est localement confluente est également confluente. Une relation  $\Rightarrow$  *termine* si il n'existe pas de chaîne de réduction infinie, c'est-à-

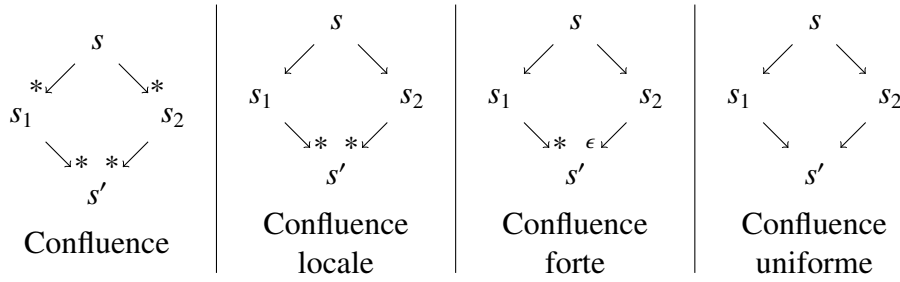
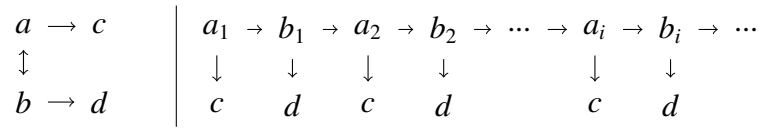
Fig 2.23: Diagrammes de confluence pour la simplification  $\rightarrow$ .

Fig 2.24: Exemples de confluence locale sans confluence globale.

dire que pour tout système  $s \in S$ , il existe un entier  $k$  tel que  $\{s' \mid s \Rightarrow^k s'\} = \emptyset$ .

**Lemme 1** (Newman, 1942). Soit  $\Rightarrow \in S \times S$  une relation bien fondée localement confluente. Alors  $\Rightarrow$  est (globalement) confluente.

Notons que sans la propriété de terminaison, la confluence locale n'implique pas la confluence globale, comme l'illustre l'exemple suivant.

**Exemple 21.** Les relations illustrées dans la Fig. 2.24 ne terminent pas, la première car elle possède une boucle, tandis que la seconde est définie sur une infinité de systèmes. Elles sont toutes les deux localement confluentes, mais ne sont pas globalement confluentes.

Premièrement, considérons l'ensemble de systèmes  $S = \{a, b, c, d\}$ , et la relation  $\Rightarrow = \{(a, b), (b, a), (a, c), (b, d)\}$ , illustrée à gauche de la Fig. 2.24. Cette relation ne termine pas, elle peut boucler infiniment entre les systèmes  $a$  et  $b$ . La relation est localement confluente. Par exemple, à partir du système  $a$ , on peut atteindre en une étape soit le système  $b$ , soit le système  $c$ . On peut alors réduire (en deux étapes)  $b$  en  $c$ , et donc atteindre le même réseau. En revanche,  $\Rightarrow$  est pas confluente. Toujours à partir de  $a$ , on peut atteindre  $c$  et  $d$ , et on ne peut ensuite jamais réduire ces réseaux en un même réseau.

De la même façon, considérons l'ensemble infini suivant  $S = \bigcup_{i>1} \{a_i, b_i\} \cup \{c, d\}$ , et la relation  $\Rightarrow = \bigcup_{i>1} \{(a_i, b_i), (b_i, a_{i+1}), (a_i, c), (b_i, d)\}$ , illustrée à droite de la Fig. 2.24. Encore une fois,  $\Rightarrow$  est localement confluente, mais elle ne termine pas et n'est donc pas confluente.

Comme notre simplification réduit toujours la taille du système, et que la relation d'ordre  $<$  est bien fondée, notre simplification termine toujours, donc on pourra utiliser la confluence locale pour prouver la confluence globale.

Enfin, on pourra considérer deux notions plus fortes de confluence (Fig. 2.23). Une simplification est *fortement confluente* si  $\Leftarrow \circ \Rightarrow \subseteq \Rightarrow^* \circ \Leftarrow^\epsilon$ . Notons que cette notion est légèrement plus faible que la notion  $\Leftarrow \circ \Rightarrow \subseteq \Rightarrow^\epsilon \circ \Leftarrow^\epsilon$ . Elle est *uniformément confluente* si  $\Leftarrow \circ \Rightarrow \subseteq \equiv \cup (\Rightarrow^1 \circ \Leftarrow^1)$ . Clairement, la confluence uniforme implique la confluence forte, qui implique la confluence locale, ainsi que la confluence globale (Huet, 1980). De plus, la confluence uniforme implique que toutes les simplifications complètes partant d'un même réseau ont la même taille (Niehren, 2000).

Afin de prouver les propriétés de confluence, on définit les deux notions suivantes. Soit  $s$  un système, et  $D_1$  et  $D_2$  deux règles de simplification applicables sur  $s$ , produisant les systèmes  $s_1$  et  $s_2$ . On dit que ces règles *ne génèrent pas de paires critiques* sur  $s$  si  $s_1 = s_2$ . On dit que *les paires critiques se ferment uniformément* si  $D_1$  est toujours applicable sur  $s_2$ , et respectivement  $D_2$  sur  $s_1$ , et qu'elles produisent alors le même réseau  $s'$  (modulo l'équivalence structurelle).





## CHAPITRE 3

---

### Congruence des attracteurs et simplification non déterministe

---

On s'intéresse tout d'abord à la simplification de réseaux de réactions tout en préservant certaines propriétés de ces réseaux, avec une sémantique non déterministe. Ces propriétés dépendent de l'objectif de la modélisation. Lors de la simplification de réseaux de Pétri (Berthelot, 1985; Lee-Kwang, Favrel, & Baptiste, 1987; Desel, 1990) par exemple, on cherche principalement à préserver des propriétés de vivacité, ou leur caractère borné. La sémantique opérationnelle des réseaux de réactions, présentée dans la Sec. 2.1, est en général trop forte. Les simplifications possibles tout en préservant le graphe d'accessibilité sont alors mineures. Il est donc nécessaire de définir une nouvelle sémantique, décrivant ce qui nous intéresse vraiment dans le comportement des réseaux.

En modélisation biologique, on s'intéresse souvent au comportement final, stable, que peut avoir un réseau de réactions. Lorsque l'on étudie la sémantique déterministe, de tels comportements correspondent à ce qu'on appelle les états d'équilibre du système, lorsque la concentration des espèces moléculaires reste constante au cours du temps. Avec une sémantique non déterministe, on peut représenter le comportement final et régulier des réseaux en utilisant la notion d'*attracteurs*. Un attracteur, comme pour la sémantique booléenne, est une composante fortement connexe terminale du graphe d'accessibilité. Cette notion est différente des états d'équilibre en déterministe, puisqu'un attracteur peut représenter une oscillation entre différents états. On peut alors étudier la possibilité de converger vers un attracteur, ou au contraire de diverger, c'est-à-dire de ne jamais atteindre de comportement régulier. Ces notions peuvent être définies modulo

la fonction d'observation, afin de s'abstraire de certaines informations non pertinentes.

Dans ce chapitre, on présente tout d'abord cette nouvelle sémantique, basée sur la convergence vers des attracteurs. On étudie alors un premier ensemble de règles de simplification, confluent, qui préserve la sémantique. On enrichit ensuite cet ensemble avec d'autres règles de simplification plus générales, mais non confluentes. On montre alors comment on peut utiliser la simplification pour comparer différents modèles d'un même système biologique, sur l'exemple des modèles de MAPK ([Markevich et al., 2004](#)). Enfin, on discute des propriétés de notre simplification, et de ces liens avec d'autres notions d'équivalences et d'autres relations de simplification dans la littérature.

Ce chapitre reprend les résultats de ([Madelaine et al., 2014](#)), et les étend en prouvant la confluence d'un sous ensemble de règles, en ajoutant des règles de simplifications plus générales, et en appliquant la simplification pour comparer différents réseaux du système MAPK.

## 3.1 Congruence des attracteurs

On définit dans cette section la sémantique non déterministe des attracteurs, basée sur la capacité des réseaux à atteindre des composantes fortement connexes terminales, ou au contraire à diverger.

### 3.1.1 Attracteurs

On s'intéresse ici aux observations des comportements finaux des réseaux de réactions, que l'on représente par des  $\Omega$ -attracteurs.

**Définition 8** (Attracteurs). Soit  $\Omega$  une fonction d'observation, un  $\Omega$ -attracteur (ou simplement un attracteur) est un ensemble  $\mathcal{A}$  non vide de réseaux de réactions vérifiant les deux propriétés suivantes :

- *fermeture* : tout réseau accessible depuis  $\mathcal{A}$  appartient à  $\mathcal{A}$ , c'est-à-dire  $\forall N \in \mathcal{A}. \{N \rightarrow^*\} \subseteq \mathcal{A}$ ,
- $\Omega$ -connectivité : toute observation d'un réseau de  $\mathcal{A}$  est accessible depuis tout réseau  $N$  de  $\mathcal{A}$ , c'est-à-dire :  $\forall N \in \mathcal{A}. \Omega(\mathcal{A}) \subseteq \Omega(\{N \rightarrow^*\})$ .

On dénote  $Attr(\Omega)$  (ou simplement  $Attr$ ) l'ensemble des  $\Omega$ -attracteurs.

Intuitivement, la propriété de fermeture signifie qu'une fois que l'on a atteint un attracteur, on va toujours rester dans cet attracteur. La propriété de connectivité quant à elle dit qu'on peut toujours parcourir l'ensemble de l'attracteur, modulo l'observation  $\Omega$ .

**Exemple 22.** Considérons le réseau suivant, avec une réaction produisant du  $A$  et une réaction dégradant du  $A$  :

$$N_0 = \{\emptyset \rightarrow A, A \rightarrow \emptyset\} \mid \emptyset.$$

On note  $N_i$  le réseau avec les mêmes réactions, et  $i$  molécules de  $A$  dans sa solution, pour tout  $i \geq 0$  :

$$N_i = \{\emptyset \rightarrow A, A \rightarrow \emptyset\} \mid iA.$$

La fonction d'observation observe le nombre de molécules de  $A$  :  $\Omega(N_i) = iA$ .

Alors l'ensemble  $\mathcal{A} = \{N_0 \rightarrow^*\}$  est un  $\Omega$ -attracteur. En effet,  $\mathcal{A}$  est fermé par définition. De plus, à partir d'un réseau  $N_i$ , on peut atteindre le réseau  $N_{i+1}$  en appliquant la réaction  $\emptyset \rightarrow A$ , ou le réseau  $N_{i-1}$  lorsque  $i > 0$ , en appliquant la réaction  $A \rightarrow \emptyset$ . Par récurrence directe, on peut atteindre à partir de  $N_i$  n'importe quel réseau  $N_j$  avec  $j \geq 0$ . D'où  $\mathcal{A} \subseteq \{N_i \rightarrow^*\}$ , et donc  $\mathcal{A}$  est  $\Omega$ -connecté. C'est donc un  $\Omega$ -attracteur.

On va maintenant énoncer quelques propriétés basiques des attracteurs.

**Proposition 2.** Soit  $\mathcal{A}$  et  $\mathcal{A}'$  deux  $\Omega$ -attracteurs. Les propriétés suivantes sont vérifiées :

1. pour tout  $N \in \mathcal{A}$ ,  $\Omega(\{N \rightarrow^*\}) = \Omega(\mathcal{A})$ ,
2. deux attracteurs se chevauchant ne sont pas observationnellement distinguables :  $\mathcal{A} \cap \mathcal{A}' \neq \emptyset$  implique  $\Omega(\mathcal{A}) = \Omega(\mathcal{A}')$ ,
3. tout sous-ensemble non vide et fermé de  $\mathcal{A}$  est aussi un attracteur, avec la même observation.

*Proof.* 1.  $\Omega(\mathcal{A}) \subseteq \Omega(\{N \rightarrow^*\})$  par  $\Omega$ -connectivité, et  $\Omega(\{N \rightarrow^*\}) \subseteq \Omega(\mathcal{A})$  découle de la propriété de fermeture  $\{N \rightarrow^*\} \subseteq \mathcal{A}$ ,

2. soit  $N \in \mathcal{A} \cap \mathcal{A}'$ , alors  $\Omega(\mathcal{A}) = \Omega(\{N \rightarrow^*\}) = \Omega(\mathcal{A}')$ ,

3. soit  $\mathcal{A}' \subseteq \mathcal{A}$  un ensemble non vide et fermé, et  $N \in \mathcal{A}'$ . Alors  $\Omega(\mathcal{A}') \subseteq \Omega(\mathcal{A}) = \Omega(\{N \rightarrow^*\})$ . Donc  $\mathcal{A}'$  est un attracteur, et  $\Omega(\mathcal{A}') = \Omega(\{N \rightarrow^*\}) = \Omega(\mathcal{A})$ .

□

### 3.1.2 Convergence et divergence

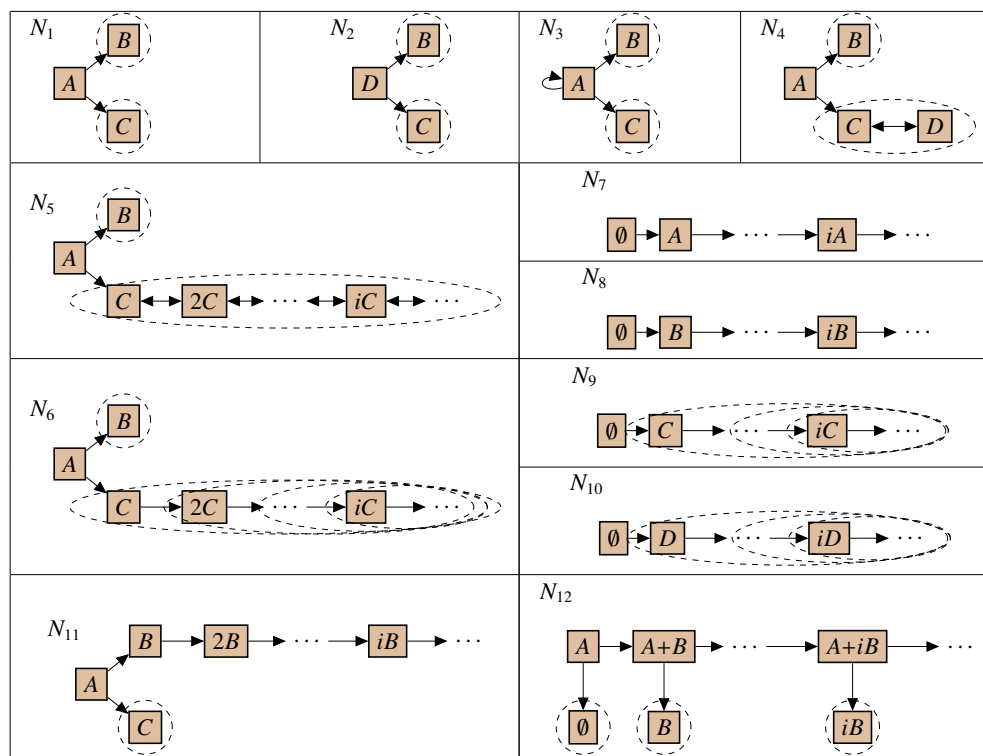
Les attracteurs représentent donc les comportements finaux des réseaux de réactions. On s'intéresse alors à la capacité d'un réseau à pouvoir converger vers ces comportements finaux, ou au contraire à pouvoir diverger.

**Définition 9** ( $\Omega$ -convergence et  $\Omega$ -divergence). Soit  $\Omega$  une fonction d'observation et  $N$  un réseau de réactions. On dit que :

- $N$  peut converger vers un ensemble d'observables  $O \subseteq \text{Observables}$ , noté  $N \downarrow_O^\Omega$ , si  $N$  peut atteindre un attracteur d'observation  $O$ , formellement :  $\exists \mathcal{A} \in \text{Attr}(\Omega). \mathcal{A} \subseteq \{N \rightarrow^*\}$  et  $\Omega(\mathcal{A}) = O$ ,
- $N$  peut diverger, noté  $N \uparrow^\Omega$ , si  $N$  peut atteindre un réseau qui ne peut pas converger :  $\exists M \in \{N \rightarrow^*\}. \forall O \subseteq \text{Observables}. M \not\downarrow_O^\Omega$ .

Notons que les ensembles d'observations  $O$  ne sont pas forcément ici, comme illustré dans l'exemple suivant.

**Exemple 23.** Le réseau  $N_0$  de l'Ex. 22 précédent peut converger vers l'ensemble  $O_1 = \{iA \mid i \geq 0\}$ , et ne peut pas diverger, donc  $N_0 \downarrow_{O_1}^\Omega$  et  $N_0 \uparrow^\Omega$ . De plus, pour tout  $O_2 \subseteq \text{Observables}$  tel que  $O_2 \neq O_1$ ,  $N_0 \not\downarrow_{O_2}^\Omega$ .



$$N_1 = \{A \rightarrow B, A \rightarrow C\} \mid A$$

$$N_2 = \{D \rightarrow B, D \rightarrow C\} \mid D$$

$$N_3 = \{A \rightarrow A, A \rightarrow B, A \rightarrow C\} \mid A$$

$$N_4 = \{A \rightarrow B, A \rightarrow C, C \rightarrow D, D \rightarrow C\} \mid A$$

$$N_5 = \{A \rightarrow B, A \rightarrow C, C \rightarrow 2C, 2C \rightarrow C\} \mid A$$

$$N_6 = \{A \rightarrow B, A \rightarrow C, C \rightarrow 2C\} \mid A$$

$$N_7 = \{\emptyset \rightarrow A\} \mid \emptyset$$

$$N_8 = \{\emptyset \rightarrow B\} \mid \emptyset$$

$$N_9 = \{\emptyset \rightarrow C\} \mid \emptyset$$

$$N_{10} = \{\emptyset \rightarrow D\} \mid \emptyset$$

$$N_{11} = \{A \rightarrow B, B \rightarrow 2B, A \rightarrow C\} \mid A$$

$$N_{12} = \{A \rightarrow A + B, A \rightarrow \emptyset\} \mid A$$

 Fig 3.1: Graphes des solutions accessibles des réseaux  $N_1$  à  $N_{12}$ .

**Exemple 24.** Considérons les réseaux  $N_1$  à  $N_{12}$ , dont les graphes des solutions accessibles sont décrits dans la Fig. 3.1. On utilise ici les espèces moléculaires  $Espèces = \{A, B, C, D, E\}$ . On suppose que la fonction d'observation permet de compter le nombre de molécules des espèces  $A$ ,  $B$  et  $E$ . Elle ne peut par contre pas distinguer les espèces  $C$  et  $D$ , et ne peut observer que l'existence de l'une de ces deux espèces, pas le nombre exacte de molécules. Formellement, on a :

$$\Omega(aA + bB + cC + dD + eE) = aA + bB + \min(c + d, 1)C + eE$$

Le réseau  $N_1 = \{A \rightarrow B, A \rightarrow C\} \mid A$  a deux attracteurs accessibles. Le premier attracteur contient uniquement le réseau  $\{A \rightarrow B, A \rightarrow C\} \mid B$ , et a donc pour observation  $\{B\}$ . Le deuxième attracteur contient le réseau  $\{A \rightarrow B, A \rightarrow C\} \mid C$ , et a pour observation  $\{C\}$ . Donc  $N_1 \downarrow_{\{B\}}^\Omega$  et  $N_1 \downarrow_{\{C\}}^\Omega$ . De plus, comme il doit converger vers l'un de ces attracteurs, il ne peut pas diverger,  $N_1 \uparrow^\Omega$ .

Le réseau  $N_2 = \{D \rightarrow B, D \rightarrow C\} \mid D$  est similaire au réseau  $N_1$ , en remplaçant l'espèce  $A$  par  $D$ . Comme cette espèce n'est utilisée que initialement, et que l'on ne s'intéresse ici qu'aux comportements finaux, ce réseau va se comporter comme le réseau  $N_1$ , c'est-à-dire  $N_2 \downarrow_{\{B\}}^\Omega$ ,  $N_2 \downarrow_{\{C\}}^\Omega$  et  $N_2 \uparrow^\Omega$ .

Le réseau  $N_3 = \{A \rightarrow A, A \rightarrow B, A \rightarrow C\} \mid A$  possède une séquence infinie de réductions  $N_3 \rightarrow N_3 \rightarrow \dots$ . Néanmoins, il ne peut pas diverger, puisque dans cette séquence il peut toujours converger. Une séquence infinie comme celle-là est parfois appelée une "divergence faible". Cette divergence n'est possible que si on suppose que la réduction n'est pas équitable, c'est-à-dire qu'on applique uniquement la réaction  $A \rightarrow A$ , en ignorant les autres, qui sont pourtant toujours applicables. On ne s'intéressera pas ici à ce genre de divergence, donc  $N_3$  se comporte comme les réseaux précédents : il peut converger vers  $\{B\}$  ou  $\{C\}$ , et ne peut pas diverger.

Le réseau  $N_4 = \{A \rightarrow B, A \rightarrow C, C \rightarrow D, D \rightarrow C\} \mid A$  a également deux attracteurs accessibles. Le premier attracteur contient encore un seul réseau d'observation  $\{B\}$ . Le deuxième contient deux réseaux,  $\{A \rightarrow B, A \rightarrow C, C \rightarrow D, D \rightarrow C\} \mid C$  et  $\{A \rightarrow B, A \rightarrow C, C \rightarrow D, D \rightarrow C\} \mid D$ . Néanmoins comme l'observation ne distingue pas  $C$  et  $D$ , cet attracteur a encore pour unique observation  $\{C\}$ . Donc, comme pour  $N_1$ , on a  $N_4 \downarrow_{\{B\}}^\Omega$ ,  $N_4 \downarrow_{\{C\}}^\Omega$ , et  $N_4 \uparrow^\Omega$ .

Le réseau suivant,  $N_5 = \{A \rightarrow B, A \rightarrow C, C \rightarrow 2C, 2C \rightarrow C\} \mid A$ , a toujours un attracteur d'observation  $\{B\}$ , et possède également un attracteur de taille infinie, contenant tous les réseaux de forme  $\{A \rightarrow B, A \rightarrow C, C \rightarrow 2C, 2C \rightarrow C\} \mid cC$  pour  $c \geq 1$ . Comme la fonction d'observation ne peut voir que la présence de l'espèce  $C$  dans la solution d'un réseau, et ne peut pas compter son nombre exacte de molécules, cet attracteur a encore pour observation  $\{C\}$ , et converge donc de la même façon que les réseaux précédents.

Le réseau  $N_6 = \{A \rightarrow B, A \rightarrow C, C \rightarrow 2C\} \mid A$ , de façon similaire au réseau  $N_5$ , peut converger vers  $\{B\}$  ou  $\{C\}$ , et ne peut pas diverger. Notons néanmoins que ce

réseau possède cette fois un nombre infini d'attracteurs, inclus les uns dans les autres, avec tous pour observation  $\{C\}$ .

Les réseaux  $N_7 = \{\emptyset \rightarrow A\} \mid \emptyset$  et  $N_8 = \{\emptyset \rightarrow B\} \mid \emptyset$  ne peuvent jamais atteindre un attracteur. Donc ils divergent tous les deux,  $N_7 \uparrow^\Omega$  et  $N_8 \uparrow^\Omega$ . En revanche, les réseaux  $N_9 = \{\emptyset \rightarrow C\} \mid \emptyset$  et  $N_{10} = \{\emptyset \rightarrow D\} \mid \emptyset$  ne divergent pas et peuvent converger vers  $\{C\}$ , puisque l'observation ne compte pas le nombre de molécules  $C$  et  $D$ .

Le réseau  $N_{11} = \{A \rightarrow B, B \rightarrow 2B, A \rightarrow C\} \mid A$  est un exemple pouvant soit converger (vers  $\{C\}$ ), soit diverger.

Enfin, le réseau  $N_{12} = R \mid A$ , avec  $R = \{A \rightarrow A+B, A \rightarrow \emptyset\}$  possède une séquence infinie de réduction  $(R \mid A) \rightarrow (R \mid A+B) \rightarrow \dots \rightarrow (R \mid A+bB) \rightarrow \dots$ . Néanmoins, il ne peut pas diverger. C'est encore une fois une divergence faible, conséquence de l'utilisation non équitable de la réaction  $A \rightarrow A+B$  sans jamais utiliser la réaction  $A \rightarrow \emptyset$ . Le réseau  $N_{12}$  peut converger vers une infinité d'attracteurs d'observations différentes:  $N_{12} \downarrow_{\{bB\}}^\Omega$  pour tout  $b \geq 0$ .

### 3.1.3 Congruence des attracteurs

On peut alors définir et étudier l'équivalence et la congruence des attracteurs : deux réseaux de réactions sont équivalents s'ils ont les mêmes possibilités de convergence et de divergence.

**Définition 10** (Équivalence des attracteurs). Soit  $\Omega$  une fonction d'observation. Deux réseaux de réactions  $N$  et  $N'$  sont  $\Omega$ -attracteurs-équivalents, noté  $N \simeq_\Omega N'$ , si :

- pour tout ensemble d'observation  $O \subseteq \text{Observables}$ ,  $N \downarrow_O^\Omega$  si et seulement si  $N' \downarrow_O^\Omega$ ,
- $N \uparrow^\Omega$  si et seulement si  $N' \uparrow^\Omega$ .

**Exemple 25.** Reprenons les réseaux de la Fig. 3.1 de l'Ex. 24. Les réseaux  $N_1$ ,  $N_2$ ,  $N_3$ ,  $N_4$ ,  $N_5$  et  $N_6$  sont tous attracteurs-équivalents, puisqu'ils peuvent tous converger soit vers  $\{B\}$ , soit vers  $\{C\}$ , et ne peuvent pas diverger.

Les réseaux  $N_7$  et  $N_8$ , même s'ils ne possèdent pas les mêmes espèces moléculaires, sont également attracteurs-équivalents. De façon plus générale, tous les réseaux qui ne peuvent pas converger, et donc peuvent uniquement diverger, sont équivalents. Ce n'est par contre pas le cas de  $N_9$  et  $N_{10}$ , qui convergent vers  $\{C\}$ . Donc  $N_7 \simeq_\Omega N_8$ , et  $N_9 \simeq_\Omega N_{10}$ . Notons qu'avec une autre fonction d'observation, on pourrait différencier d'autres réseaux. Par exemple, considérons la fonction d'observation :

$$\Omega'(aA + cC + s) = aA + cC \text{ où } A, C \notin s.$$



Elle ne compte que le nombre de  $A$  et de  $C$ , on aurait alors  $N_7 \simeq_{\Omega} N_9$ , puisqu'ils divergent tous les deux, et  $N_8 \simeq_{\Omega} N_{10}$ , puisqu'ils convergent tous les deux vers l'observation vide  $\{O\}$ .

Le réseau  $N_{11}$  est le seul réseau qui peut à la fois converger et diverger, et n'est donc pas attracteurs-équivalent aux autres réseaux. De même,  $N_{12}$  est le seul pouvant converger vers une infinité d'ensembles d'observations différents, et n'est donc pas équivalent aux autres.

L'équivalence des attracteurs est suffisante pour exprimer d'autres prédicats de convergence. Par exemple, on dit qu'un réseau  $N$  *doit converger vers*  $O$ , dénoté  $N \Downarrow_O^{\Omega}$ , si tous les réseaux atteignables peuvent converger vers  $O$ , c'est-à-dire  $\forall N' \in \{N \rightarrow^*\}. N' \Downarrow_O^{\Omega}$ . Un réseau *peut converger*,  $N \Downarrow^{\Omega}$ , s'il existe  $O$  tel qu'il puisse converger vers  $O$ . Il *doit converger*,  $N \Downarrow^{\Omega}$ , s'il ne peut pas diverger. Enfin, un réseau *doit diverger*, dénoté  $N \Uparrow^{\Omega}$ , si tous les réseaux atteignables peuvent diverger, c'est-à-dire  $\forall N' \in \{N \rightarrow^*\}. N' \Uparrow^{\Omega}$ . C'est équivalent à dire que  $N$  ne peut pas converger.

**Exemple 26.** Dans l'Ex. 24, les réseaux  $N_1$  à  $N_6$  peuvent et doivent converger. Les réseaux  $N_7$  et  $N_8$  doivent diverger. Les réseaux  $N_9$  et  $N_{10}$  doivent converger vers  $\{C\}$ .

**Proposition 3.** Si  $N \simeq_{\Omega} N'$ , alors :

- $N \Downarrow_O^{\Omega}$  si et seulement si  $N' \Downarrow_O^{\Omega}$ , pour tout  $O \subseteq \text{Observables}$ ,
- $N \Downarrow^{\Omega}$  si et seulement si  $N' \Downarrow^{\Omega}$ ,
- $N \Downarrow^{\Omega}$  si et seulement si  $N' \Downarrow^{\Omega}$ ,
- $N \Uparrow^{\Omega}$  si et seulement si  $N' \Uparrow^{\Omega}$ .

*Proof.* Supposons que  $N \simeq_{\Omega} N'$ , et soit  $O \subseteq \text{Observables}$ . On va d'abord montrer que  $N \Downarrow_O^{\Omega}$  si et seulement si d'une part  $\neg N \Uparrow^{\Omega}$  et d'autre part  $\forall O' \subseteq \text{Observables}, N \Downarrow_{O'}^{\Omega}$  implique  $O = O'$ . Supposons que  $N \Downarrow_{O'}^{\Omega}$  et que  $N$  puisse diverger, alors il existe un  $N_1 \in \{N \rightarrow^*\}$  qui ne peut pas converger, donc en particulier qui ne peut pas converger vers  $O$ , ce qui contredit  $N \Downarrow_O^{\Omega}$ . Supposons que  $N \Downarrow_O^{\Omega}$  et que  $N$  puisse aussi converger vers  $O' \neq O$ , alors il existe un attracteur  $\mathcal{A}$  tel que  $\Omega(\mathcal{A}) = O'$ , et un réseau  $N_1 \in \{N \rightarrow^*\} \cap \mathcal{A}$ . Alors  $N_1$  ne peut pas converger vers  $O$ , ce qui est contradictoire. Réciproquement, si  $\neg N \Uparrow^{\Omega}$  et que  $\forall O' \subseteq \text{Observables}, N \Downarrow_{O'}^{\Omega}$  implique  $O = O'$ . La première partie implique que pour tout  $N_1 \in \{N \rightarrow^*\}$ , il existe  $O'$  tel que  $N_1 \Downarrow_{O'}^{\Omega}$ , et donc  $N \Downarrow_{O'}^{\Omega}$ . La deuxième partie implique  $O = O'$ , et donc  $N \Downarrow_O^{\Omega}$ . On a alors les résultats suivants :

- si  $N \Downarrow_O^{\Omega}$ , alors  $\neg N \Uparrow^{\Omega}$  et  $\forall O' \subseteq \text{Observables}, N \Downarrow_{O'}^{\Omega}$  implique  $O = O'$ . Comme  $N \simeq_{\Omega} N'$ , cela implique la même chose pour  $N'$ , et donc  $N' \Downarrow_O^{\Omega}$ .

- $N \downarrow^\Omega$  si et seulement si  $\exists O. N \downarrow_O^\Omega$ , si et seulement si  $\exists O. N' \downarrow_O^\Omega$ , si et seulement si  $N' \downarrow^\Omega$ .
- $N \Downarrow^\Omega$  si et seulement si  $N \Uparrow^\Omega$  si et seulement si  $N' \Downarrow^\Omega$ .
- $N \Uparrow^\Omega$  si et seulement si  $N \Downarrow^\Omega$  si et seulement si  $N' \Uparrow^\Omega$ .

□

Enfin, on ne veut pas considérer des réseaux isolés, mais plutôt en tant que sous-modules d'un réseau plus grand. Cela permet de définir des règles de simplification plus simples, puisque définies localement et non globalement. Les contextes permettent aussi d'appliquer la simplification à des sous-réseaux sans connaître exactement le réseau global. On va donc étendre notre équivalence en une relation de congruence.

**Définition 11** (Congruence des attracteurs). Soit  $\Omega$  une fonction d'observation et  $\mathcal{I}$  un sous-ensemble d'espèces observables. Deux réseaux de réactions  $N$  et  $N'$  sont  $\Omega, \mathcal{I}$ -attracteurs-congruents, noté  $N \sim_{\Omega, \mathcal{I}} N'$ , si pour tout réseau  $N''$  tel que  $\text{Espèces}(N'') \cap \mathcal{I} = \emptyset$ , on a  $N \mid N'' \simeq_\Omega N' \mid N''$ .

### 3.1.4 Liens avec le graphe des solutions accessibles

Dans cette section, on prouve que deux réseaux ayant le même graphe des solutions accessibles sont attracteurs-équivalents. Cette propriété nous aidera notamment plus tard pour prouver que nos simplifications préservent la congruence des attracteurs. On rappelle que  $\mathcal{G}_{sol}(N)$  est le graphe des solutions accessibles de  $N$ , et  $\overline{\mathcal{G}_{sol}(N)}$  le graphe obtenu en supprimant les boucles.

**Lemme 2.** Soit  $\Omega$  une fonction d'observation, et  $N$  et  $N'$  deux réseaux de réactions, tels que  $\overline{\mathcal{G}_{sol}(N)} = \overline{\mathcal{G}_{sol}(N')}$ . Alors pour toute fonction d'observation  $\Omega$ , on a  $N \simeq_\Omega N'$ .

*Proof.* Soit  $N = R \mid s$  et  $N' = R' \mid s'$  deux réseaux, tels que  $\overline{\mathcal{G}_{sol}(N)} = \overline{\mathcal{G}_{sol}(N')}$ . Ces graphes doivent avoir la même racine, donc  $s = s'$ .

On va montrer que pour toute solution  $s_1, s_2$  apparaissant dans le graphe des solutions accessibles  $\overline{\mathcal{G}_{sol}(N)}$ , on a :

$$R \mid s_1 \rightarrow R \mid s_2 \quad \text{si et seulement si} \quad R' \mid s_1 \rightarrow R' \mid s_2. \quad (3.1)$$

Ce résultat découle directement de l'égalité  $\overline{\mathcal{G}_{sol}(N)} = \overline{\mathcal{G}_{sol}(N')}$ .

Pour tout réseau  $N_1 = R \mid s_1$ , on définit alors la fonction  $f(N_1) = R' \mid s_1$ . Soit  $\mathcal{A} \subseteq \{N \rightarrow^*\}$  un  $\Omega$ -attracteur accessible depuis  $N$ , on va montrer que  $\mathcal{A}' = f(\mathcal{A})$  est également un  $\Omega$ -attracteur, avec la même observation, et accessible depuis  $N'$ .

Il existe  $N_1 \in \mathcal{A}$  avec  $N_1 \in \{N \rightarrow^*\}$ , donc par 3.1,  $f(N_1) \in \mathcal{A}'$  vérifie  $f(N_1) \in \{N' \rightarrow^*\}$ , donc  $f(\mathcal{A}')$  est accessible. Soit  $N'_1 \in \mathcal{A}'$  et  $N'_2 \in \{N'_1 \rightarrow^*\}$ . Il existe  $N_1 = f^{-1}(N'_1) \in \mathcal{A}$ , et par 3.1,  $N_2 = f^{-1}(N'_2) \in \{N_1 \rightarrow^*\}$ . Comme  $\mathcal{A}$  est fermé,  $N_2 \in \mathcal{A}$ , et donc  $N'_2 \in \mathcal{A}'$ . Donc  $\mathcal{A}'$  est fermé. Soit  $o \in \Omega(\mathcal{A}')$ , et  $N'_1 \in \mathcal{A}'$ . Alors il existe  $N'_2 \in \mathcal{A}'$  avec  $\Omega(N'_2) = o$ . Mais alors  $N_1 = f^{-1}(N'_1) \in \mathcal{A}$  et de même  $N_2 = f^{-1}(N'_2) \in \mathcal{A}$ . Donc  $o = \Omega(N_2) \in \Omega(\mathcal{A})$ , et par  $\Omega$ -connectivité,  $o \in \Omega(\{N_1 \rightarrow^*\})$ . Donc il existe  $N_3 \in \{N_1 \rightarrow^*\}$  avec  $\Omega(N_3) = o$ . Mais on a alors  $f(N_3) \in \{N'_1 \rightarrow^*\}$  avec  $\Omega(f(N_3)) = o$ , donc  $o \in \Omega(\{N'_1 \rightarrow^*\})$ , donc  $\Omega(\mathcal{A}') \subseteq \Omega(\{N'_1 \rightarrow^*\})$ . Donc  $\mathcal{A}'$  est  $\Omega$ -connecté, donc c'est un  $\Omega$ -attracteur, et il a trivialement la même observation que  $\mathcal{A}$ . Par symétrie, pour tout  $\Omega$ -attracteur  $\mathcal{A}'$  accessible depuis  $N'$ ,  $f^{-1}(\mathcal{A}')$  est un  $\Omega$ -attracteur accessible depuis  $N$ . Donc pour tout ensemble d'observables  $O \subseteq \text{Observables}$ ,  $N \downarrow_O^\Omega$  si et seulement si  $N' \downarrow_O^\Omega$ . De plus,  $N \uparrow^\Omega$  si et seulement si il existe  $N_1 \in \{N \rightarrow^*\}$  tel que  $\forall O \subseteq \text{Observables}$ ,  $N_1 \not\downarrow_O^\Omega$ . Mais avec le même raisonnement que ci-dessus,  $N_1 \not\downarrow_O^\Omega$  si et seulement si  $f(N_1) \not\downarrow_O^\Omega$ . Donc  $N \uparrow^\Omega$  si et seulement si  $N' \uparrow^\Omega$ . Et donc  $N$  et  $N'$  sont attracteurs-équivalents.  $\square$

Par exemple, dans la Fig. 3.1, les réseaux  $N_1$  et  $N_3$  ont le même graphe réduit, et sont donc équivalents pour toute fonction d'observation.

La réciproque est trivialement fautive, comme on peut par exemple le voir avec les réseaux  $N_1$  et  $N_2$  sur la Fig. 3.1, qui sont attracteurs-équivalents, mais n'ont pas le même graphe des solutions accessibles.

## 3.2 Ensemble confluent de règles de simplification

Dans la section précédente, on a défini une nouvelle notion de congruence, basée sur les attracteurs et les capacités de convergence et divergence des réseaux. On présente maintenant un premier ensemble de règles de simplification, correcte pour la congruence des attracteurs. Ces règles sont suffisantes pour simplifier le réseau Tet-On (Fig. 2.7). De plus, on montre que cette simplification est confluente.

On suppose dans la suite que l'observation  $\Omega$  est fixée, ainsi que l'ensemble des espèces internes  $\mathcal{I}$ . La simplification est alors notée  $\Rightarrow$ .

### 3.2.1 Règles de simplification

Les cinq règles de simplification, ainsi qu'une règle supplémentaire pour rendre la simplification contextuelle, sont décrites ci-dessous et présentées formellement dans la Fig. 3.2.

**Réaction Cyclique** La première règle de simplification, (**RÉACTION CYCLIQUE**), supprime une réaction qui, lorsqu'elle est appliquée, ne modifie pas le réseau.

**Activateur Présent** La règle de simplification (**ACTIVATEUR PRÉSENT**) supprime un activateur  $A$  d'une réaction  $s_1 + aA \rightarrow s_2 + aA$ , s'il y a toujours au moins  $a$  molécules de  $A$  dans la solution du réseau initial et de tous les réseaux accessibles. L'espèce  $A$  doit donc être interne, et les autres réactions ne peuvent pas diminuer le nombre de molécules de  $A$ .

**Espèce Inutile** La règle (**ESPÈCE INUTILE**) supprime une espèce  $A$ , si elle est interne, non observable, et inutile. Une espèce  $A$  est inutile si, pour toute réaction  $r$ , soit  $A$  n'est pas un réactant de  $r$  et donc n'est pas nécessaire pour appliquer la réaction ( $\forall s_1 \rightarrow s_2 \in R. A \notin s_1$ ), soit  $r$  ne peut modifier uniquement la quantité de molécules de  $A$ , pas celles des autres espèces ( $\forall r \in R'. \forall B \neq A. stoic_r(B) = 0$ ). On peut alors complètement supprimer l'espèce  $A$  du réseau.

**Réaction Inversible** La règle (**RÉACTION INVERSIBLE**) supprime deux réactions qui sont l'inverse l'une de l'autre :  $s \rightarrow A$  et  $A \rightarrow s$ , avec  $A \notin s$ . La règle remplace également toutes occurrences de  $A$  par la solution  $s$ . L'espèce  $A$  doit être interne, et doit avoir la même observation que  $s$ . Les deux autres conditions sont nécessaires pour obtenir une simplification confluente. La condition  $|s| > 1$  garantie que  $s$  est une vraie solution. Dans le cas contraire, on pourrait avoir  $s = B$ , et la règle pourrait simplifier  $A$  ou  $B$  de façon non confluente. Notons qu'on pourrait

(RÉACTION CYCLIQUE)

$$\frac{}{\{s \rightarrow s\} \mid \emptyset \Rightarrow^{Cyc} \emptyset \mid \emptyset}$$

(ACTIVATEUR PRÉSENT)

$$\frac{\begin{array}{l} A \in \mathcal{I} \quad a_0 \geq a > 0 \\ A \notin s_1 \quad \forall \mathbf{r} \in \mathbf{R}. \text{stoic}_{\mathbf{r}}(A) \geq 0 \end{array}}{\{s_1 + aA \rightarrow s_2 + aA\} \cup \mathbf{R} \mid a_0A \Rightarrow^{Act} \{s_1 \rightarrow s_2\} \cup \mathbf{R} \mid a_0A}$$

(ESPÈCE INUTILE)

$$\frac{\begin{array}{l} A \in \mathcal{I} \quad \Omega(A) = 0 \quad \forall s_1 \rightarrow s_2 \in \mathbf{R}. A \notin s_1 \\ \forall \mathbf{r} \in \mathbf{R}'. \forall B \neq A. \text{stoic}_{\mathbf{r}}(B) = 0 \end{array}}{\mathbf{R} \cup \mathbf{R}' \mid a_0A \Rightarrow^{Inu} \mathbf{R} \setminus A \mid \emptyset}$$

(RÉACTION INVERSIBLE)

$$\frac{\begin{array}{l} A \in \mathcal{I} \quad \Omega(A) = \Omega(s) \quad A \notin s \\ |s| > 1 \quad \forall \mathbf{r} \in \mathbf{R}. \text{stoic}_{\mathbf{r}}(A) = 0 \end{array}}{\mathbf{R} \cup \{s \rightarrow A, A \rightarrow s\} \mid a_0A \Rightarrow^{Rev} \mathbf{R}[A := s] \mid a_0s}$$

(CASCADE)

$$\frac{\begin{array}{l} A \in \mathcal{I} \quad \Omega(A) = 0 \\ \forall \mathbf{r} \in \mathbf{R}. \exists a > 0. \exists s. \mathbf{r} = s \rightarrow s + aA \wedge A \notin s \\ \forall \mathbf{r}' \in \mathbf{R}'. \exists a' > 0. \exists s'. \mathbf{r}' = a'A \rightarrow s' \end{array}}{\mathbf{R} \cup \mathbf{R}' \cup \{A \rightarrow \emptyset\} \mid \emptyset \Rightarrow^{Cas} \mathbf{R} \diamond_A \mathbf{R}' \mid \emptyset}$$

(CONTEXTE)

$$\frac{\text{Espèces}(N'') \cap \mathcal{I} = \emptyset \quad N \Rightarrow^D N'}{N \mid N'' \Rightarrow^{\bar{D}} N' \mid N''}$$

Fig 3.2: Ensemble confluent de règles de simplification pour la congruence des attracteurs, avec la fonction d'observation  $\Omega$  et les espèces internes  $\mathcal{I}$ .

être confluent sans cette condition en ajoutant une notion d'isomorphisme entre réseaux, qui permettrait de renommer une espèce en une autre, par exemple ici  $A$  en  $B$  ou réciproquement. L'autre condition impose que  $A$  ne puisse apparaître dans d'autres réactions uniquement en tant qu'activateur, et donc en particulier permet d'éviter le cas où  $A$  est dans deux paires de réactions inversibles différentes, avec des solutions  $s$  et  $s'$ .

**Cascade** Enfin, la règle (**CASCADE**) simplifie une cascade de réactions. Le réseau initial contient un ensemble de réactions  $R$ , de forme  $s \rightarrow s + aA$ , qui peuvent produire des molécules de  $A$  en présence d'activateurs  $s$ . Les molécules de  $A$  peuvent être dégradées par une réaction  $A \rightarrow \emptyset$ , ou peuvent être utilisées par les réactions de  $R'$ , de forme  $a'A \rightarrow s'$ . L'espèce  $A$  doit être interne, non observable, et non présente dans la solution du réseau initial. On la supprime alors complètement, et on combine les réactions pour produire directement les solutions  $s'$  en présence des activateurs  $s$ . On définit la *combinaison modulo  $A$*  de deux réactions  $r = s_1 \rightarrow s_2$  et  $r' = s'_1 \rightarrow s'_2$  par :

$$r \diamond_A r' = (s_1 + s'_1) \setminus A \rightarrow (s_2 + s'_2) \setminus A.$$

On étend l'opérateur aux ensembles de réactions :

$$R \diamond_A R' = \{r \diamond_A r' \mid r \in R, r' \in R'\}.$$

En particulier, notons que si  $R$  ou  $R'$  est vide, on a alors  $R \diamond_A R' = \emptyset$ . Dans la règle (**CASCADE**), si  $R$  est vide, alors l'espèce  $A$  ne peut jamais être produite, donc les réactions de  $R'$  ne sont pas utilisables. Si  $R'$  est vide, alors  $A$  est une espèce inutile.

**Contexte** La règle (**CONTEXTE**) étend la simplification aux contextes, et permet donc d'appliquer nos règles de simplification de façon locale sur les sous-réseaux, tout en restant valable pour le réseau global. Si on peut simplifier  $N$  en  $N'$  avec une règle  $D$ , alors on peut simplifier  $N \mid N''$  en  $N' \mid N''$  avec la règle  $\tilde{D}$ . Notons qu'on a trivialement  $\tilde{\tilde{D}} = \tilde{D}$ .

On définit la relation  $\Rightarrow$  comme l'union des règles de simplification de la Fig. 3.2. Remarquons que si  $N \Rightarrow N'$ , alors il existe  $D \in \{Cyc, Act, Inu, Rev, Cas\}$  tel que  $N \Rightarrow^{\tilde{D}} N'$  (avec éventuellement un contexte vide).

### 3.2.2 Propriétés

On montre que la relation  $\Rightarrow$  est bien une simplification, c'est-à-dire qu'elle réduit la taille du réseau. On prouve ensuite qu'elle est correcte pour la congruence  $\sim_\Omega$ , et qu'elle est confluite. On étudie alors la complexité de nos règles de simplification.

**Simplification** Montrons tout d'abord que nos règles réduisent bien la taille des réseaux.

**Définition 12** (Taille et ordre des réseaux). Soit  $N = \mathbb{R} \mid s_0$  un réseau. Sa *taille*  $|N|$  est le triplet  $(n, m, p)$ , avec :

- $n = \text{Espèces}(N)$  le nombre d'espèces moléculaires de  $N$ ,
- $m = |\mathbb{R}|$  le nombre de réactions de  $N$ ,
- $p = |s_0| + \sum_{s_1 \rightarrow s_2 \in \mathbb{R}} |s_1 + s_2|$  la taille combinée de toutes les solutions du réseau.

Un réseau  $N$  de taille  $(n, m, p)$  est alors *plus petit* qu'un autre réseau  $N'$  de taille  $(n', m', p')$ , noté  $N < N'$ , si  $(n, m, p) < (n', m', p')$  avec l'ordre lexicographique, c'est-à-dire :

- $n < n'$ ,
- ou  $n = n'$  et  $m < m'$ ,
- ou  $n = n'$ ,  $m = m'$  et  $p < p'$ .

**Exemple 27.** Par exemple, pour Michaelis-Menten, le réseau  $N_{\text{MM}}$  (Fig. 2.5) possède 4 espèces ( $S$ ,  $E$ ,  $C$  et  $P$ ), 3 réactions, et chaque réaction est de taille 3 ( $r_1$  consomme 2 molécules et en produit 1, tandis que  $r_2$  et  $r_3$  consomment 1 molécule et en produisent 2). Donc la taille du réseau est  $|N_{\text{MM}}| = (4, 3, 9)$ .

**Lemme 3.** L'ordre  $<$  sur les réseaux de réactions est bien fondé.

**Proposition 4.** Si  $N \Rightarrow N'$ , alors  $N' < N$ .

*Proof.* Soit  $(n, m, p)$  la taille de  $N$ , et  $(n', m', p')$  la taille de  $N'$ . Si  $N \Rightarrow^{\bar{D}} N'$  avec  $D \in \{\text{Inu}, \text{Rev}, \text{Cas}\}$ , alors  $n' = n - 1 < n$ . Si  $D = \text{Cyc}$ , alors  $n = n'$  et  $m' = m - 1 < m$ . Si  $D = \text{Act}$ , alors  $n = n'$ ,  $m = m'$  et  $p' = p - 2a < p$ .  $\square$

**Exactitude** On va maintenant prouver qu'un réseau simplifié est attracteurs-congruent au réseau initial.

**Théorème 1.** La simplification  $\Rightarrow_{\Omega, \mathcal{I}}$  est correcte pour la congruence des attracteurs  $\sim_{\Omega, \mathcal{I}}$ .

*Proof.* On doit donc prouver que si  $N \Rightarrow_{\Omega, \mathcal{I}} N'$ , alors  $N \sim_{\Omega, \mathcal{I}} N'$ , c'est-à-dire que pour tout contexte  $N''$  tel que  $\text{Espèces}(N'') \cap \mathcal{I} = \emptyset$ , on a  $(N \mid N'') \simeq_{\Omega, \mathcal{I}} (N' \mid N'')$ . On note  $R''$  les réactions de  $N''$ . Les preuves sont en général basées sur l'existence de fonctions  $f$  et  $g$  reliant les réseaux initiaux et les réseaux simplifiés. La preuve de la règle (**CASCADE**) est un peu plus compliquée, et nécessite de réordonner certaines séquences de réactions.

Pour la règle (**RÉACTION CYCLIQUE**), on a trivialement le même graphe réduit des solutions accessibles  $\overline{\mathcal{G}_{sol}(N)} = \overline{\mathcal{G}_{sol}(N')}$ , donc  $\overline{\mathcal{G}_{sol}(N \mid N'')} = \overline{\mathcal{G}_{sol}(N' \mid N'')}$ . Donc, d'après le Lem. 2,  $N \mid N''$  et  $N' \mid N''$  sont attracteurs-équivalents, et donc  $N \sim_{\Omega, \mathcal{I}} N'$ .

Considérons la règle (**ACTIVATEUR PRÉSENT**), appliqué sur un activateur  $A$ . Le réseau  $N$  contient une réaction  $s_1 + aA \rightarrow s_2 + aA$ , ainsi qu'un ensemble de réactions ne pouvant qu'augmenter le nombre de molécules de  $A$ . Comme  $A$  est une espèce interne, les réactions  $R''$  du contexte ne peuvent pas modifier la quantité de  $A$ . Donc, par récurrence immédiate, il y a toujours au moins  $a_0$  molécules de  $A$  dans tout réseau accessible depuis  $N \mid N''$  ou  $N' \mid N''$ . Soit  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  un réseau accessible depuis le réseau initial dans son contexte. Il existe alors une solution  $s''$  avec  $A \notin s''$  et un entier  $a_1 \geq a_0 \geq a$  tels que :

$$N_1 = \{s_1 + aA \rightarrow s_2 + aA\} \cup R \cup R'' \mid a_1A + s''.$$

Pour tout réseau de cette forme, on note  $f$  la fonction définie par :

$$f(N_1) = \{s_1 \rightarrow s_2\} \cup R \cup R'' \mid a_1A + s''.$$

La fonction  $f$  est alors une bijection entre les réseaux de  $\{(N \mid N'') \rightarrow^*\}$  et ceux de  $\{(N' \mid N'') \rightarrow^*\}$ . On a alors la propriété suivante :

$$N_1 \rightarrow N_2 \quad \text{si et seulement si} \quad f(N_1) \rightarrow f(N_2).$$

En effet, si la réaction utilisée pour transformer  $N_1$  en  $N_2$  appartient à  $R$  ou à  $R''$ , alors elle existe aussi dans  $f(N_1)$  (et inversement). Si la réaction utilisée est  $s_1 + aA \rightarrow s_2 + aA$ , alors on peut utiliser  $s_1 \rightarrow s_2$  à la place. Et inversement, si on utilise la réaction  $s_1 \rightarrow s_2$  pour transformer  $f(N_1)$  en  $f(N_2)$ , alors comme  $a_1 \geq a$ , on peut utiliser  $s_1 + aA \rightarrow s_2 + aA$  pour transformer  $N_1$  en  $N_2$ .

Par récurrence immédiate,  $N_1 \rightarrow^* N_2$  si et seulement si  $f(N_1) \rightarrow^* f(N_2)$ . On a alors  $f(\{N_1 \rightarrow^*\}) = \{f(N_1) \rightarrow^*\}$ . De plus,  $f(N) = N'$ . Et tout ensemble de réseaux  $\mathcal{A}$  est fermé et accessible depuis  $N \mid N''$  si et seulement si  $f(\mathcal{A})$  est fermé et accessible depuis  $N' \mid N''$ .



La fonction  $f$  préserve aussi trivialement l'observation :  $\Omega(f(N_1)) = \Omega(N_1)$ . Supposons alors qu'on a un attracteur  $\mathcal{A}$ . Par définition, il est  $\Omega$ -connecté, c'est-à-dire que pour tout  $N_1 \in \mathcal{A}$ , on a  $\Omega(\mathcal{A}) \subseteq \Omega(\{N_1 \rightarrow^*\})$ . Alors, pour tout réseau  $f(N_1) \in f(\mathcal{A})$ , on aura :

$$\begin{aligned} \Omega(f(\mathcal{A})) &= \Omega(\mathcal{A}) \\ &\subseteq \Omega(\{N_1 \rightarrow^*\}) \\ &\subseteq \Omega(f(\{N_1 \rightarrow^*\})) \\ &\subseteq \Omega(\{f(N_1) \rightarrow^*\}). \end{aligned}$$

Donc  $f(\mathcal{A})$  est aussi  $\Omega$ -connecté. La réciproque est identique. Donc  $\mathcal{A}$  est un  $\Omega$ -attracteur si et seulement si  $f(\mathcal{A})$  est un  $\Omega$ -attracteur. Donc pour tout  $O \in \text{Observables}$ , on a  $(N \mid N'') \downarrow_O^\Omega$  si et seulement si  $(N' \mid N'') \downarrow_O^\Omega$ . De plus,  $(N \mid N'') \uparrow^\Omega$  si et seulement si il existe  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  tel que pour tout ensemble d'observations  $O \subseteq \text{Observables}$ ,  $N_1 \not\downarrow_O^\Omega$ . Mais alors, avec le même raisonnement que ci-dessus,  $f(N_1) \in \{(N' \mid N'') \rightarrow^*\}$  vérifie  $\forall O \subseteq \text{Observables}$ ,  $f(N_1) \not\downarrow_O^\Omega$ , et donc  $(N' \mid N'') \uparrow^\Omega$ , et réciproquement. Donc  $N \sim_\Omega N'$ .

Considérons la règle de simplification (**ESPÈCE INUTILE**), appliquée sur une espèce  $A$ . On note  $R$  les réactions  $s_1 \rightarrow s_2$  telles que  $A \notin s_1$ , et  $R'$  les réactions ne pouvant modifier que la quantité de  $A$ , comme définies dans la règle. Tout réseau accessible  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  est de forme  $N_1 = R \cup R' \cup R'' \mid a_1 A + s''$ , avec  $s''$  une solution telle que  $A \notin s''$  et  $a_1$  un entier. Soit  $f$  la fonction définie pour tout réseau  $N_1$  de cette forme par  $f(N_1) = R \setminus A \cup R'' \mid s''$ . Alors pour toute paire de réseaux accessibles  $N_1, N_2 \in \{(N \mid N'') \rightarrow^*\}$ , on va montrer que :

$$N_1 \rightarrow N_2 \quad \text{implique} \quad f(N_1) \rightarrow^\epsilon f(N_2).$$

C'est en fait immédiat : supposons que  $N_1 \rightarrow_r N_2$ . Si  $r \in R''$ , alors  $A \notin r$ , donc  $f(N_1) \rightarrow_r f(N_2)$ . Si  $r \in R$ , alors  $r \setminus A \in R \setminus A$  est applicable, et on obtient bien le réseau  $f(N_2)$ . Si  $r \in R'$ , alors  $f(N_1) = f(N_2)$ . Par récurrence immédiate, on a :

$$N_1 \rightarrow^* N_2 \quad \text{implique} \quad f(N_1) \rightarrow^* f(N_2). \quad (3.2)$$

Réciproquement, tout réseau  $N'_1 \in \{(N' \mid N'') \rightarrow^*\}$  est de forme  $N'_1 = R \setminus A \cup R'' \mid s''$  avec  $A \notin s''$ . On définit alors l'ensemble des réseaux  $g(N'_1) = \{(R \setminus A \cup R' \cup R'' \mid a'_1 A + s'') \mid a'_1 \in \mathbb{N}\}$ . Soit  $N'_1, N'_2 \in \{(N' \mid N'') \rightarrow^*\}$ , on va montrer qu'on a :

$$N'_1 \rightarrow N'_2 \quad \text{implique} \quad \forall N_1 \in g(N'_1). \exists N_2 \in g(N'_2). N_1 \rightarrow N_2.$$

Supposons en effet que  $N'_1 \rightarrow_{r'} N'_2$ . Si  $r' \in R''$ , alors  $r'$  est directement applicable dans  $N_1$ , et la propriété est vraie. Si  $r' \in R \setminus A$ , alors il existe  $r \in R$  telle que

$r' = r \setminus A$ . Or, d'après les conditions de la règle de simplification, on sait qu'il n'y a pas de molécules de  $A$  dans les réactants de  $r$ . Donc  $r$  est toujours applicable dans  $N_1$ , et produit un réseau  $N_2 \in g(N_2')$ . Par récurrence immédiate, pour tout réseau  $N_1', N_2' \in \{(N' \mid N'') \rightarrow^*\}$ , on a :

$$N_1' \rightarrow^* N_2' \quad \text{implique} \quad \forall N_1 \in g(N_1'). \exists N_2 \in g(N_2'). N_1 \rightarrow^* N_2. \quad (3.3)$$

Remarquons de plus que  $\Omega(f(N_1)) = \Omega(N_1)$ , et que pour tout  $N_1 \in g(N_1')$ ,  $\Omega(N_1) = \Omega(N_1')$ . De plus  $N_1 \in g(f(N_1))$ ,  $f(g(N_1')) = \{N_1'\}$ ,  $f(N) = N'$  et  $N \in g(N')$ .

Soit  $\mathcal{A} \subseteq \{(N \mid N'') \rightarrow^*\}$  un  $\Omega$ -attracteur accessible depuis le réseau initial dans son contexte. Soit  $N_1 \in \mathcal{A}$ ,  $N_1' = f(N_1)$  et  $\mathcal{A}' = \{N_1' \rightarrow^*\}$ . On va montrer que  $\mathcal{A}'$  est également un  $\Omega$ -attracteur, accessible depuis  $N' \mid N''$ . Il est fermé par définition, et accessible d'après 3.2. Soit  $N_2' \in \mathcal{A}'$  et  $o \in \Omega(\mathcal{A}')$  une observation. Il existe donc  $N_3' \in \mathcal{A}'$  tel que  $o = \Omega(N_3')$ . On a  $N_1 \in g(N_1')$ , donc il existe  $N_2 \in g(N_2') \cap \mathcal{A}$  d'après 3.3. De même, il existe  $N_3 \in g(N_3') \cap \mathcal{A}$ , avec  $\Omega(N_3) = o$ . Donc  $o \in \Omega(\mathcal{A})$ . Comme  $\mathcal{A}$  est  $\Omega$ -connecté, il existe  $N_4 \in \{N_2 \rightarrow^*\}$  avec  $o = \Omega(N_4)$ . Par la propriété 3.2, on a alors  $f(N_4) \in \{N_2' \rightarrow^*\}$ , donc  $o \in \Omega(\{N_2' \rightarrow^*\})$ . Donc  $\mathcal{A}'$  est  $\Omega$ -connecté, et c'est un attracteur.

Réciproquement, si  $\mathcal{A}'$  est un attracteur accessible depuis  $N' \mid N''$ , et  $N_1' \in \mathcal{A}'$ , alors d'après 3.3 il existe  $N_1 \in g(N_1') \cap \{(N \mid N'') \rightarrow^*\}$ . On va montrer que  $\mathcal{A} = \{N_1 \rightarrow^*\}$  est bien un attracteur. Il est trivialement fermé, et accessible. Soit  $N_2 \in \mathcal{A}$  et  $o \in \Omega(\mathcal{A})$ , c'est-à-dire qu'il existe  $N_3 \in \mathcal{A}$  avec  $o = \Omega(N_3)$ . Par 3.2,  $N_2' = f(N_2)$  et  $N_3' = f(N_3)$  appartiennent à  $\mathcal{A}'$ . Donc  $o \in \Omega(\mathcal{A}')$ , donc par  $\Omega$ -connectivité il existe  $N_4' \in \{N_2' \rightarrow^*\}$  avec  $o = \Omega(N_4')$ . Mais alors par 3.3 il existe  $N_4 \in g(N_4') \cap \{N_2 \rightarrow^*\}$ , et on a donc  $o \in \Omega(\{N_2 \rightarrow^*\})$ . Donc  $\mathcal{A}$  est un attracteur.

Trivialement, les attracteurs  $\mathcal{A}$  et  $\mathcal{A}'$  ont la même observation, donc  $(N \mid N'') \downarrow_o^\Omega$  si et seulement si  $(N' \mid N'') \downarrow_o^\Omega$ . La preuve pour la divergence est similaire. Donc les réseaux  $N$  et  $N'$  sont attracteurs-congruents.

Considérons maintenant la règle (**RÉACTION INVERSIBLE**). On note  $r_1 = s \rightarrow A$  et  $r_2 = A \rightarrow s$  les réactions supprimées par la règle de simplification. Soit  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  un réseau accessible depuis le réseau initial dans son contexte. Soit  $R''$  les réactions de  $N''$ . Il existe une solution  $s_1$  telle que :

$$N_1 = \{r_1, r_2\} \cup R \cup R'' \mid s_1.$$

Pour tout réseau de cette forme, on dénote par  $f$  la fonction définie par :

$$f(N_1) = R[A := s] \cup R'' \mid s_1[A := s].$$

On va alors montrer que :

$$N_1 \rightarrow N_2 \quad \text{implique} \quad f(N_1) \rightarrow^\epsilon f(N_2).$$

Soit  $r$  la réaction utilisée, c'est-à-dire  $N_1 \rightarrow_r N_2$ . Si  $r \in R$ , alors  $r' = r[A := s] \in R[A := s]$  et  $f(N_1) \rightarrow_{r'} f(N_2)$ . Si  $r = s \rightarrow A$  ou  $r = A \rightarrow s$ , alors  $f(N_1) = f(N_2)$ . Enfin, si  $r \in R''$ , alors  $A \notin r$ , et donc  $r$  est aussi applicable dans  $f(N_1)$ , donc  $f(N_1) \rightarrow_r f(N_2)$ . Par récurrence directe, on a :

$$N_1 \rightarrow^* N_2 \quad \text{implique} \quad f(N_1) \rightarrow^* f(N_2). \quad (3.4)$$

Réciproquement, soit  $N'_1 \in \{(N' \mid N'') \rightarrow^*\}$  un réseau accessible depuis le réseau simplifié dans son contexte. Il existe  $s''$  avec  $A \notin s''$  telle que :

$$N'_1 = R[A := s] \cup R'' \mid s''.$$

Pour tout réseau de cette forme, on dénote par  $g$  la fonction définie par :

$$g(N'_1) = \{s \rightarrow A, A \rightarrow s\} \cup R \cup R'' \mid s''.$$

On montre alors que :

$$N'_1 \rightarrow N'_2 \quad \text{implique} \quad g(N'_1) \rightarrow^* g(N'_2).$$

En effet, supposons que  $N'_1 \rightarrow_{r'} N'_2$ , avec  $r'$  la réaction utilisée. Si  $r' \in R''$ , on peut directement appliquer  $r'$ , donc  $g(N'_1) \rightarrow_{r'} g(N'_2)$ . Si  $r' \in R[A := s]$ , il existe une réaction  $r \in R$  telle que  $r' = r[A := s]$ . À partir de  $g(N'_1)$ , on peut alors transformer certaines occurrences de la solution  $s$  en molécules de  $A$  avec la réaction  $s \rightarrow A$ , puis utiliser la réaction  $r$ , et enfin retransformer les molécules de  $A$  restantes en  $s$  avec  $A \rightarrow s$ . On a donc  $g(N'_1) \rightarrow_{r_1}^* \rightarrow_r \rightarrow_{r_2}^* g(N'_2)$ . Par récurrence directe, on en déduit que :

$$N'_1 \rightarrow^* N'_2 \quad \text{implique} \quad g(N'_1) \rightarrow^* g(N'_2). \quad (3.5)$$

Remarquons aussi que :

$$f(g(N'_1)) = N'_1 \quad \text{et} \quad g(f(N_1)) \rightarrow^* N_1 \quad \text{et} \quad N_1 \rightarrow^* g(f(N_1)). \quad (3.6)$$

La partie gauche est immédiate. Pour la partie droite, il suffit d'utiliser les réactions  $r_1$  et  $r_2$  pour transformer des molécules de  $A$  en  $s$  ou réciproquement. Enfin, notons que les fonctions  $f$  et  $g$  trivialement préservent l'observation, puisque  $\Omega(A) = \Omega(s)$ , que  $f(N) = N'$  et que  $N \rightarrow^* g(N')$  en utilisant la réaction  $s \rightarrow A$ .

Soit  $\mathcal{A} \subseteq \{(N \mid N'') \rightarrow^*\}$  un attracteur accessible depuis le réseau initial, et  $N_1 \in \mathcal{A}$ . On va montrer que  $\mathcal{A}' = \{f(N_1) \rightarrow^*\}$  est aussi un attracteur, avec la même observation. Il est fermé par définition. Soit  $N'_2 \in \mathcal{A}'$  un réseau, et  $o \in \Omega(\mathcal{A}')$  une observation. Donc il existe un réseau  $N'_3 \in \mathcal{A}'$  avec  $\Omega(N'_3) = o$ . Par la propriété 3.5,  $f(N_1) \rightarrow^* N'_3$  implique  $g(f(N_1)) \rightarrow^* g(N'_3)$ , et donc  $N_1 \rightarrow^* g(N'_3)$  par 3.6. Donc  $g(N'_3) \in \{N_1 \rightarrow^*\} \subseteq \mathcal{A}$  et  $o = \Omega(N'_3) = \Omega(g(N'_3)) \in \Omega(\mathcal{A})$ . De même,

$g(N'_2) \in \mathcal{A}$ . Donc la propriété de  $\Omega$ -connectivité implique que  $o \in \Omega(\{g(N'_2) \rightarrow^*\})$ , c'est-à-dire qu'il existe  $N_4 \in \{g(N'_2) \rightarrow^*\}$  avec  $\Omega(N_4) = o$ . Donc par 3.4,  $f(N_4) \in \{f(g(N'_2)) \rightarrow^*\} = \{N'_2 \rightarrow^*\}$ . Donc  $o \in \Omega(\{N'_2 \rightarrow^*\})$ , donc  $\mathcal{A}'$  est  $\Omega$ -connecté, donc c'est un attracteur. Et trivialement  $\Omega(\mathcal{A}') = \Omega(\mathcal{A})$ .

La réciproque est exactement similaire, et donc  $(N \mid N'') \downarrow_O^\Omega$  si et seulement si  $(N' \mid N'') \downarrow_O^\Omega$ .

Pour la divergence, si  $(N \mid N'') \uparrow^\Omega$ , alors il existe  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  tel que pour tout  $O \subseteq \text{Observables}$ ,  $N_1 \downarrow_O^\Omega$ . Mais, avec le même raisonnement qu'avant, cela implique que  $f(N_1) \downarrow_O^\Omega$ , et donc que  $(N' \mid N'') \uparrow^\Omega$ . Par symétrie,  $(N \mid N'') \uparrow^\Omega$  si et seulement si  $(N' \mid N'') \uparrow^\Omega$ .

Donc  $N$  et  $N'$  sont attracteurs-congruents.

Enfin, considérons la règle (**CASCADE**). On note  $r_d$  la réaction de dégradation  $A \rightarrow \emptyset$ ,  $R$  les réactions produisant  $A$  et  $R'$  celles consommant  $A$ , comme définies dans la règle. Pour tout réseau  $N_1$  accessible depuis  $N \mid N''$ , il existe  $s''$  avec  $A \notin s''$  et un entier  $a$  tels que :

$$N_1 = R \cup R' \cup \{r_d\} \cup R'' \mid aA + s''.$$

Pour tout réseau de la forme ci-dessus, on définit la fonction  $f$  par :

$$f(N_1) = (R \diamond_A R') \cup R'' \mid s''.$$

Soit  $k$  un entier, et  $N_0, N_k \in \{(N \mid N'') \rightarrow^*\}$  deux réseaux accessibles, tels que  $A$  ne soit pas dans les solutions de ces deux réseaux, c'est-à-dire  $A \notin s(N_0), s(N_k)$ . On veut montrer que :

$$N_0 \rightarrow^* N_k \quad \text{implique} \quad f(N_0) \rightarrow^* f(N_k). \quad (3.7)$$

Supposons que  $N_0 \rightarrow^k N_k$ , décomposons la séquence de réduction en  $k$  étapes :

$$N_0 \rightarrow N_1 \rightarrow \dots \rightarrow N_{k-1} \rightarrow N_k.$$

On suppose que pour tout  $0 < i < k$ ,  $A \in s(N_i)$ . En effet, pour toute séquence de forme  $N_i \rightarrow^* N_j$ , si il n'y a jamais aucune molécule de  $A$  dans les solutions initiales des réseaux  $N_i$  à  $N_j$ , alors c'est que seules les réactions du contexte  $M_1$  sont utilisées, et on a alors directement  $f(N_i) \rightarrow^* f(N_j)$ . Si la séquence possède des réseaux avec des molécules de  $A$  et des réseaux sans molécules de  $A$ , prouver la propriété 3.7 sur les sous-séquences possédant du  $A$  est suffisant.

Premièrement, remarquons que la réaction  $r_d$  commute à droite, c'est-à-dire que pour toute réaction  $r$  et toute paire de réseaux  $N_i, N_{i+2}$  telle que  $N_i \rightarrow_{r_d} \rightarrow_r N_{i+2}$ , la réduction  $N_i \rightarrow_r \rightarrow_{r_d} N_{i+2}$  est aussi possible. On réécrit alors la séquence précédente en plaçant toutes les occurrences de  $r_d$  à la fin de la séquence :

$$N_0 \rightarrow N_1 \rightarrow \dots \rightarrow N_{k'} \xrightarrow{r_d} N_k,$$

tel que  $r_d$  ne soit jamais utilisée dans la séquence  $N_0 \rightarrow^* N_{k'}$ . Notons qu'on a directement  $f(N_{k'}) = f(N_k)$ , et donc  $f(N_{k'}) \rightarrow^* f(N_k)$ . On s'intéresse maintenant uniquement à la séquence  $N_0 \rightarrow^* N_{k'}$ .

On a  $A \notin s(N_0)$  mais  $A \in s(N_1)$ . Donc la réaction  $r_1$  utilisée pour la réduction  $N_0 \rightarrow_{r_1} N_1$  appartient à  $R$ . On va partitionner la séquence  $N_0 \rightarrow^* N_{k'}$ , afin de faire apparaître les applications des réactions de  $R$ . On a donc :

$$\begin{aligned} N_0 &\rightarrow_{r_1} N_{1,0} \xrightarrow{R' \cup R''}^{k_1} N_{1,k_1} \\ &\rightarrow_{r_2} N_{2,0} \xrightarrow{R' \cup R''}^{k_2} N_{2,k_2} \\ &\dots \\ &\rightarrow_{r_j} N_{j,0} \xrightarrow{R' \cup R''}^{k_j} N_{j,k_j}, \end{aligned}$$

tel que  $N_{j,k_j} = N_{k'}$ , et que pour tout  $1 \leq i \leq j$ ,  $r_i \in R$ . On utilise donc une réaction de  $R$ , suivie d'une séquence de réactions de  $R'$  et de  $R''$ , puis une nouvelle réaction de  $R$ , etc. Notons que les réactions  $r_i$  ne sont pas forcément distinctes. Pour tout  $1 \leq i \leq j$ , la réaction  $r_i \in R$  est de forme  $r_i = s_i \rightarrow s_i + a_i A$ , avec  $a_i > 0$ , donc  $s_i \subseteq s(N_{i,0})$ . On a directement  $f(N_0) = f(N_{1,0})$  et pour tout  $1 \leq i < j$ ,  $f(N_{i,k_i}) = f(N_{i+1,0})$ . Il suffit donc de prouver la propriété 3.7 pour toute séquence  $N_{i,0} \xrightarrow{R' \cup R''}^{k_i} N_{i,k_i}$ .

Dans cette sous-séquence, seules des réactions de  $R'$  ou de  $R''$  peuvent être utilisées. Or, les réactions de  $R''$  n'ont pas de molécules de  $A$ . De plus, les réactions de  $R'$ , si l'on s'abstrait de l'espèce  $A$ , ne peuvent que créer de nouvelles molécules. Cela implique qu'une réaction  $r' \in R$  va commuter à gauche avec une réaction  $r'' \in R''$ . En d'autres termes, si  $N_{i,i'} \xrightarrow{r''} N_{i,i'+2}$ , alors  $N_{i,i'} \xrightarrow{r'} \rightarrow_{r''} N_{i,i'+2}$ . On peut donc déplacer toutes les réactions de  $R'$  vers la gauche, et obtenir une séquence de forme :

$$N_{i,0} \rightarrow_{r'_0} N_{i,1} \rightarrow_{r'_1} \dots \rightarrow_{r'_{k'_i-1}} N_{i,k'_i} \xrightarrow{R''} N_{i,k_i}$$

avec pour tout  $0 \leq i' < k'_i$ ,  $r'_{i'} = a'_{i'} A \rightarrow s'_{i'} \in R'$ . Puisque la séquence de réseaux  $N_{i,k'_i} \xrightarrow{R(M_1)}^* N_{i,k_i}$  n'utilise que des réactions du contexte  $M$ , sans molécule de  $A$ , cela directement implique  $f(N_{i,k'_i}) \xrightarrow{R''}^* f(N_{i,k_i})$ . De plus, pour tout  $0 \leq i' < k'_i$ , on a  $s_i \in N_{i,i'}$ . En effet, on a vu plus haut que  $s_i \in N_{i,0}$ , et comme  $A \notin s_i$ , les réactions  $r'_{i'}$  ne modifient pas la présence de  $s_i$  dans les réseaux. Mais, par définition de (CASCADE), on peut alors appliquer la réaction  $r_i \diamond_A r'_{i'}$  sur  $f(N_{i,i'})$ , et obtenir  $f(N_{i,i'+1})$ . On a donc :

$$f(N_{i,1}) \rightarrow_{r_1 \diamond r'_1} f(N_{i,2}) \rightarrow_{r_1 \diamond r'_2} \dots \rightarrow_{r_1 \diamond r'_{k'_i-1}} f(N_{i,k'_i}) \xrightarrow{R''}^* f(N_{i,k_i}).$$

De même, on a donc finalement  $f(N_0) \rightarrow^* f(N_k)$ , donc la propriété 3.7 est vérifiée.

Réciproquement, pour tout réseau  $N'_1$  accessible depuis  $N' \mid N''$ , il existe  $s''$  avec  $A \notin s''$  telle que :

$$N'_1 = R \diamond_A R' \cup R'' \mid s''.$$

Pour tout réseau de cette forme, on définit la fonction  $g$  par :

$$g(N'_1) = R \cup R' \cup \{r_d\} \cup R'' \mid s''.$$

On montre que pour tout réseau  $N'_0, N'_k \in \{(N' \mid N'') \rightarrow^*\}$ , on a :

$$N'_0 \rightarrow^* N'_k \quad \text{implique} \quad g(N'_0) \rightarrow^* g(N'_k). \quad (3.8)$$

La preuve est directe par récurrence. Si  $N'_0 \rightarrow_{R'} N'_k$ , alors  $g(N'_0) \rightarrow_{R'} g(N'_k)$ . Si  $N'_0 \rightarrow_{r \circ_A r'} N'_k$ , avec  $r = s \rightarrow s + aA$  et  $r' = a'A \rightarrow s'$ , alors la séquence suivante est correcte :  $g(N'_0) \rightarrow_r^* \rightarrow_{r'}^* \rightarrow_{r_d}^* g(N'_k)$ . Donc 3.8 est vérifiée.

On va maintenant prouver que pour tout ensemble d'observables  $O$ , on a :

$$(N \mid N'') \downarrow_O^\Omega \quad \text{si et seulement si} \quad (N' \mid N'') \downarrow_O^\Omega.$$

Premièrement, remarquons que comme  $\Omega(A) = \emptyset$ , les deux fonctions précédentes préservent l'observation :  $\Omega(f(N_1)) = \Omega(N_1)$ , et  $\Omega(g(N'_1)) = \Omega(N'_1)$ . De plus, on a  $N'_1 = f(g(N'_1))$ . Si  $A \notin s(N_1)$ , alors  $N_1 = g(f(N_1))$ . Sinon,  $N_1 \rightarrow_{r_d}^* g(f(N_1))$ .

Soit  $\mathcal{A}$  un  $\Omega$ -attracteur accessible depuis  $N \mid N''$ , et  $N_1 \in \mathcal{A}$  un réseau appartenant à l'attracteur. On a alors  $(N \mid N'') \rightarrow_{r_d}^* N_1 \rightarrow_{r_d}^* N_1 \setminus A$ . Comme  $A \notin s(N \mid N'')$ , la propriété 3.7 implique que  $N'_1 = f(N_1 \setminus A)$  est accessible depuis  $N' \mid N''$ . On va montrer que  $\mathcal{A}' = \{N'_1 \rightarrow^*\}$  est un  $\Omega$ -attracteur. Il est fermé par définition. Soit  $N'_2 \in \mathcal{A}'$ , et  $o \in \Omega(\mathcal{A}')$ . Il existe  $N'_3 \in \mathcal{A}'$  avec  $o = \Omega(N'_3)$ . Donc  $N'_1 \rightarrow^* N'_3$ , et  $g(N'_1) \rightarrow^* g(N'_3)$  par 3.8. Or  $g(N'_1) = N_1 \setminus A \in \mathcal{A}$ , donc  $o = g(N'_3) \in \Omega(\mathcal{A})$ . De même,  $g(N'_2) \in \mathcal{A}$ . Par  $\Omega$ -connectivité, on a donc  $o \in \Omega(\{g(N'_2) \rightarrow^*\})$ , donc il existe  $N_4$  tel que  $o = \Omega(N_4) = \Omega(N_4 \setminus A)$  et  $g(N'_2) \rightarrow_{r_d}^* N_4 \rightarrow_{r_d}^* N_4 \setminus A$ . Or  $A \notin g(N'_2)$ , donc  $N'_2 = f(g(N'_2)) \rightarrow^* f(N_4 \setminus A)$ . Donc  $o \in \{N'_2 \rightarrow^*\}$ , et  $\mathcal{A}'$  est  $\Omega$ -connecté. Donc c'est un attracteur, et trivialement  $\Omega(\mathcal{A}) = \Omega(\mathcal{A}')$ .

La réciproque, ainsi que la preuve pour la divergence, fonctionnent de la même manière. On a donc bien  $N \sim_\Omega N'$ .

Enfin, considérons la règle (CONTEXTE), avec  $N \Rightarrow^{\bar{D}} N'$ , pour une règle  $D \in \{Cyc, Act, Inu, Rev, Cas\}$ . Il existe  $N_1, N'_1$  et  $N''_1$  avec  $N = N_1 \mid N''_1$  et  $N' = N'_1 \mid N''_1$ , tels que  $N_1 \Rightarrow^D N'_1$ . On a donc par induction  $N_1 \sim_\Omega N'_1$ , et donc  $N \sim_\Omega N'$ .  $\square$

**Confluence** On va maintenant montrer que cet ensemble de règles de simplification est confluant. Il n'y a pas besoin de "confluence modulo", l'équivalence structurelle étant directement l'égalité des réseaux.

**Théorème 2** (Confluence). La relation de simplification  $\Rightarrow$  est confluente pour  $(Nets, =)$ .

*Proof.* On va montrer la confluence locale, c'est-à-dire que pour toute paire de règles  $D_1, D_2 \in \{Cyc, Act, Inu, Rev, Cas\}$ , pour tout réseau  $N \in Nets$ , si  $N \Rightarrow^{D_1} N_1$  et  $N \Rightarrow^{D_2} N_2$ , alors il existe un réseau  $N' \in Nets$  tel que  $N_i \Rightarrow^* N'$ . Comme la relation  $\Rightarrow$  termine toujours, la confluence locale implique directement la confluence globale. La plupart des paires sont directes. La paire  $D_1 = D_2 = Cas$ , appliquée sur deux espèces différentes, est la plus compliquée, et nécessite d'étudier dans le détail les réactions des différents réseaux simplifiés.

Commençons par la règle  $D_1 = Cyc$ , appliquée sur une réaction  $r_1 = s \rightarrow s$ . Si  $D_2 = Cyc$ , appliquée sur  $r_2$ , alors soit  $r_1 = r_2$ , et les deux règles ne génèrent pas de paires critiques, c'est-à-dire  $N_1 = N_2$ . Soit  $r_1 \neq r_2$ , et les paires critiques se ferment uniformément, c'est-à-dire  $N_1 \Rightarrow^{D_2} N'$  et  $N_2 \Rightarrow^{D_1} N'$ .

Supposons  $D_2 = Act$ , appliquée sur un activateur  $A_2$  et une réaction  $r_2 = s_1 + a_2 A_2 \rightarrow s_2 + a_2 A_2$ . Si  $r_1 = r_2$ , et donc  $s_1 = s_2$ , alors l'application de  $D_1$  supprime la réaction, tandis que l'application de  $D_2$  la remplace par  $r'_2 = s_1 \rightarrow s_1$ . Mais on peut alors utiliser (**RÉACTION CYCLIQUE**) sur  $r'_2$  et la supprimer, on obtient donc le même réseau. Si  $r_1 \neq r_2$ , les paires critiques se ferment uniformément.

Si  $D_2 = Inu$ , appliquée sur une espèce  $A_2$ . On note  $R_2$  les réactions n'ayant pas de molécules de  $A_2$  en réactants, et  $R'_2$  celles ne pouvant modifier que la quantité de  $A_2$ , comme définies dans la règle (**ESPÈCE INUTILE**). Si  $A_2 \in r_1$ , alors  $r_1 \in R'_2$ . Donc on peut soit appliquer  $D_1$  puis  $D_2$ , soit appliquer directement  $D_2$ , dans les deux cas on obtient le même réseau. Si  $A_2 \notin r_1$ , les paires critiques se ferment uniformément.

Si  $D_2 = Rev$ , appliquée sur une espèce  $A_2$  et des réactions  $r_2 = s_2 \rightarrow A_2$  et  $r'_2 = A_2 \rightarrow s_2$ . Comme  $A_2 \notin s_2$ , trivialement  $r_1 \neq r_2$  et  $r_1 \neq r'_2$ , donc les paires critiques se ferment uniformément.

Enfin, si  $D_2 = Cas$ , appliquée sur  $A_2$ . On note  $R_2$  les réactions produisant  $A_2$ ,  $r_2 = A_2 \rightarrow \emptyset$  la réaction de dégradation, et  $R'_2$  les autres réactions avec  $A_2$  en réactant, comme dans (**CASCADE**). On a  $r_1 \neq r_2$ , et  $r_1 \notin R_2$ . Supposons que  $r_1 \in R'_2$ , c'est-à-dire  $r_1 = a' A_2 \rightarrow a' A_2$  avec  $a' > 0$ . Dans ce cas, pour toute réaction  $r = s_2 \rightarrow s_2 + a_2 A_2 \in R_2$ , la règle  $D_2$  crée la réaction  $r \diamond_{A_2} r_1 = s_2 \rightarrow s_2$ . On peut alors supprimer toutes ces réactions en utilisant  $|R|$  fois la règle (**RÉACTION CYCLIQUE**). On obtient alors le même réseau que si on avait commencé par  $D_1$  suivie de  $S_2$ . Si  $r_1 \notin R'_2$ , les paires critiques se ferment uniformément.

Supposons ensuite que  $D_1 = Act$ , appliquée sur un activateur  $A_1$  et une réaction  $r_1 = s_1 + a_1 A_1 \rightarrow s_2 + a_1 A_1$ . Supposons que  $D_2 = Act$ , appliquée sur un activateur  $A_2$  et une réaction  $r_2 = s'_1 + a_2 A_2 \rightarrow s'_2 + a_2 A_2$ . Si  $A_1 = A_2$  et  $r_1 = r_2$ ,

alors  $a_1 = a_2$  (par la condition  $A_1 \notin s_1$ ), et  $D_1$  et  $D_2$  ne génèrent pas de paires critiques sur  $N$ . Dans tous les autres cas, les paires critiques se ferment uniformément.

Si  $D_2 = Inu$ , appliquée sur une espèce inutile  $A_2$ . On note  $R_2$  les réactions n'ayant pas de molécules de  $A_2$  en réactants, et  $R'_2$  celles ne pouvant modifier que la quantité de  $A_2$ , comme définies dans la règle (**ESPÈCE INUTILE**). Si  $r_1 \in R'_2$ , alors  $r_1 = s + A_1A \rightarrow s + a_2A$ , avec  $a_1 \leq a_2$ . Après application de  $D_1$ , la réaction réduite  $r'_1 = s \rightarrow s$  appartient toujours à  $R'_2$ , et est donc supprimée par  $D_2$ . Si on applique  $D_2$  en premier,  $r_1$  est directement supprimée sans avoir à appliquer  $D_1$ . Si  $r_1 \notin R'_2$ , alors  $A_1 \neq A_2$ , et les paires critiques se ferment uniformément.

Si  $D_2 = Rev$ , appliquée sur une espèce  $A_2$  et des réactions  $r_2 = s \rightarrow A_2$  et  $r'_2 = A_2 \rightarrow s$ . La condition  $A_2 \notin s$  dans  $D_2$  implique  $r_1 \neq r_2$  et  $r_1 \neq r'_2$ . De plus, la condition  $stoic_{r'_2}(A_1) \geq 0$  dans (**ACTIVATEUR PRÉSENT**) implique  $A_1 \neq A_2$ , tandis que  $stoic_{r_2}(A_1) \geq 0$  implique alors  $A_2 \notin s$ . Donc l'application de  $D_1$  ne modifie pas les réactions  $r_2$  et  $r'_2$ . Et inversement on peut toujours appliquer  $D_1$  après  $D_2$ . Donc les paires critiques se ferment uniformément.

Si  $D_2 = Cas$ , appliquée sur  $A_2$ . On note  $R_2$  les réactions produisant  $A_2$ ,  $r_2 = A_2 \rightarrow \emptyset$  la réaction de dégradation, et  $R'_2$  les autres réactions avec  $A_2$  en réactant, comme dans (**CASCADE**). Trivialement  $r_1 \neq r_2$ . La condition  $stoic_{r_2}(A_1) \geq 0$  implique  $A_1 \neq A_2$ . Ce qui implique alors  $r_1 \notin R'_2$ . Supposons que  $r_1 \in R_2$ , c'est-à-dire qu'il existe  $a_2 > 0$  et  $s$  tel que  $r_1 = s + a_1A_1 \rightarrow s + a_1A_1 + a_2A_2$ . Si on applique d'abord  $D_1$ ,  $r_1$  est remplacée par  $r'_1 = s \rightarrow s + a_2A_2$ . On peut alors toujours appliquer  $D_2$  sur  $(R_2 \setminus r_1) \cup \{r'_1\}$ ,  $R'_2$  et  $r_2$ , et on obtient les réactions  $((R_2 \setminus r_1) \diamond_{A_2} R'_2) \cup (\{r'_1\} \diamond_{A_2} R'_2)$ . Si on commence par la règle  $D_2$ , on obtient les réactions  $R_2 \diamond_{A_2} R'_2$ . En particulier, les réactions de  $\{r_1\} \diamond_{A_2} R'_2$  sont de forme  $s + a_1A_1 \rightarrow s + a_1A_1 + s'$ , et peuvent toutes être simplifiées en appliquant récursivement  $D_1$ , pour obtenir des réactions de forme  $s \rightarrow s + s'$ . Or ces réactions sont en fait exactement les réactions de  $\{r'_1\} \diamond_{A_2} R'_2$ . On obtient donc le même réseau final. Si  $r_1 \notin R_2$ , alors les paires critiques se ferment uniformément.

Considérons maintenant le cas de la règle  $D_1 = Inu$ , appliquée sur  $A_1$ . On note  $R_1$  les réactions n'ayant pas de molécules de  $A_1$  en réactants, et  $R'_1$  celles ne pouvant modifier que la quantité de  $A_1$ , comme définies dans la règle (**ESPÈCE INUTILE**). Supposons que  $D_2 = Inu$ , appliquée sur  $A_2$ , avec  $R_2$  les réactions sans  $A_2$  en réactants et  $R'_2$  celles ne modifiant que  $A_2$ . Si  $A_1 = A_2 = A$ , remarquons que les ensembles de réactions peuvent être différents, tant que  $R_1 \cup R'_1 = R_2 \cup R'_2$  et que les conditions pour la simplification sont respectées. En effet une réaction de forme  $s \rightarrow s + aA$ , avec  $A \notin s$ , vérifie aussi bien les conditions de l'ensemble  $R$  que de l'ensemble  $R'$ . Si on applique  $D_1$ , on obtient le réseau  $R_1 \setminus A$ . Or, on peut remarquer que les réactions de  $(R_1 \cap R'_2) \setminus A$  sont des réactions cycliques, pouvant être supprimées par la règle (**RÉACTION CYCLIQUE**). On obtient alors le réseau  $(R_1 \cap$



$R_2 \setminus A$ . Par symétrie, on obtient le même réseau si on applique d'abord  $D_2$ , puis **(RÉACTION CYCLIQUE)**. Si  $A_1 \neq A_2$ , les paires critiques se ferment uniformément.

Supposons que  $D_2 = Rev$ , appliquée sur une espèce  $A_2$  et des réactions  $r_2 = s \rightarrow A_2$  et  $r'_2 = A_2 \rightarrow s$ . Si  $A_1 = A_2 = A$ , alors  $r'_2 \notin R_1$ , donc  $r'_2 \in R'_1$ . Donc pour tout  $B \neq A$ ,  $s(B) = 0$ . De plus,  $A \notin s$ , donc  $s = \emptyset$ , et  $r_2 = \emptyset \rightarrow A \in R_1$ . L'application de  $D_1$  produit le réseau  $R_1 \setminus A$ , tandis que  $D_2$  produit le réseau  $R_1[A := \emptyset] \cup R'_1[A := \emptyset] = (R_1 \setminus A) \cup (R'_1 \setminus A)$ . Or, pour toute réaction  $r \in (R'_1 \setminus A)$ , pour tout  $B \neq A$ ,  $stoic_r(B) = 0$ , et  $A \notin r$ , donc  $r$  est une réaction cyclique. On peut donc appliquer récursivement la règle **(RÉACTION CYCLIQUE)**, et obtenir le réseau  $R_1 \setminus A$ . Si  $A_1 \neq A_2$ , les paires critiques se ferment uniformément.

Si  $D_2 = Cas$ , appliquée sur  $A_2$ . On note  $R_2$  les réactions produisant  $A_2$ , la réaction de dégradation  $r_2 = A_2 \rightarrow \emptyset$ , et  $R'_2$  les autres réactions avec  $A_2$  en réactant, comme dans **(CASCADE)**. Supposons que  $A_1 = A_2 = A$ . Dans ce cas,  $R_1 \subseteq R_2$ , et  $R'_2 \subseteq R'_1$ . Plus précisément, le réseau initial contient les réactions suivantes :

- la réaction de dégradation  $r_2 = A \rightarrow \emptyset$ ,
- des réactions produisant du  $A$  :  $R_2 = \{s \rightarrow s + aA \mid A \notin s\}$ . Ces réactions peuvent appartenir à  $R_1$  ou à  $R'_1$ ,
- l'ensemble des autres réactions, qui ne modifient que la quantité de  $A$  :  $R'_2 = \{a'A \rightarrow a''A \mid a', a'' > 0\} \subseteq R'_1$ .

Si on applique  $D_1$ , on obtient alors le réseau  $R_1 \setminus A \subseteq R_2 \setminus A$ . Ces réactions sont toutes de forme  $s \rightarrow s$ , et peuvent être supprimées en appliquant la règle **(RÉACTION CYCLIQUE)**. Si on applique  $D_2$  en première, on obtient le réseau  $R_2 \diamond_A R'_2$ . Mais on obtient encore une fois des réactions de forme  $s \rightarrow s$ , pouvant être supprimées avec **(RÉACTION CYCLIQUE)**. On obtient donc bien le même réseau. Si  $A_1 \neq A_2$ , les conditions de  $D_1$  impliquent  $A_1 \notin R_2$ . Supposons qu'il existe  $r' \in R'_2$  avec  $A_1 \in r'$ . Donc il existe  $a_1, a_2 > 0$ ,  $s'$  tel que  $r' = a_2 A_2 \rightarrow s' + a_1 A_1$ , et  $A_1 \notin s'$ , et donc  $r' \in R_1$ . En appliquant  $D_1$ ,  $r'$  est transformée en  $r'' = a_2 A_2 \rightarrow s'$ . On peut donc toujours appliquer  $D_2$ , et obtenir des réactions  $R_2 \diamond_{A_2} r''$  de forme  $s \rightarrow s + (s' \setminus A)$ . Si on commence par  $D_2$ , on obtient des réactions  $R_2 \diamond_{A_2} r'$  de forme  $s \rightarrow s + (s' \setminus A_2) + a_1 A_1$ . L'espèce  $A_1$  est toujours inutile, donc on peut appliquer  $D_1$ , et on obtient bien les réactions de  $R_2 \diamond_{A_2} r''$ . Si  $A_1 \neq A_2$  et  $A_1 \notin R'_2$ , les paires critiques se ferment uniformément.

Supposons maintenant que  $D_1 = Rev$ , appliquée sur une espèce  $A_1$  et des réactions  $r_1 = s_1 \rightarrow A_1$  et  $r'_1 = A_1 \rightarrow s_1$ . Soit  $D_2 = Rev$ , appliquée sur  $A_2$  et des réactions  $r_2 = s_2 \rightarrow A_2$  et  $r'_2 = A_2 \rightarrow s_2$ . Si  $A_1 = A_2$ , alors  $r_1 = r_2$ ,  $r'_1 = r'_2$  et  $s_1 = s_2$ , sinon la condition  $stoic_{r_2}(A_1) = 0$  (resp.  $r'_1$ ) de  $D_1$  contredit la condition  $A_2 \notin s_2$  de  $D_2$ . Donc dans ce cas les deux règles ne génèrent pas de

paires critiques. Sinon, la condition  $stoic_{r_2}(A_1) = 0$  implique soit  $s_2 = A_1$  (et  $s_1 = A_1$ ), soit  $A_1 \notin s_2$  (et  $A_2 \notin s_1$ ). Le premier cas n'est pas possible, à cause des conditions  $|s_1| > 1$  et  $|s_2| > 1$ . Dans le second cas, les paires critiques se ferment uniformément.

Si la règle est  $D_2 = Cas$ , appliquée à une espèce  $A_2$ . On note  $R_2$  les réactions produisant  $A_2$ ,  $r_2 = A_2 \rightarrow \emptyset$  la réaction de dégradation, et  $R'_2$  les autres réactions avec  $A_2$  en réactant, comme dans (CASCADE). Supposons que  $A_1 = A_2 = A$ . Dans ce cas, les différentes conditions imposent  $r_2 = r'_1 = A \rightarrow \emptyset$ ,  $s = \emptyset$ ,  $r_1 = \emptyset \rightarrow A$ ,  $R_2 = \{r_1\}$ , et pour tout  $r' \in R'_2$ ,  $stoic_{r'}(A) = 0$ . L'application de  $D_1$  produit alors les réactions  $R'_2 \setminus A$ . L'application de  $D_2$  produit  $\emptyset \rightarrow A \diamond_A R'_2 = R'_2 \setminus A$ . Donc dans ce cas là, les deux règles ne génèrent pas de paires critiques. Si  $A_1 \neq A_2$ , on ne peut pas avoir  $A_2 \in s_1$ . En effet, il faudrait alors que  $r'_1 \in R_2$ , donc que  $A_1 \in s'_1$ , ce qui contredit une condition de  $D_1$ . Donc les paires critiques se ferment uniformément.

Enfin, si  $D_1 = D_2 = Cas$ , avec  $D_i$  appliquée sur  $A_i$ . Si  $A_1 = A_2$ , les deux règles ne génèrent pas de paires critiques. Sinon, on note  $R_1$  les réactions produisant  $A_1$ ,  $r_1 = A_1 \rightarrow \emptyset$  la réaction de dégradation, et  $R'_1$  les autres réactions avec  $A_1$  en réactant, et symétriquement pour  $A_2$ . On a alors  $R_1 \cap R_2 = \emptyset$  et  $R'_1 \cap R'_2 = \emptyset$ . Plus précisément, les réactions du réseau initial peuvent être partitionner de la façon suivante :

- les réactions de dégradation  $r_1 = A_1 \rightarrow \emptyset$  et  $r_2 = A_2 \rightarrow \emptyset$ ,
- les réactions produisant  $A_1$ , avec  $A_2$  en réactant :  $R_1 \cap R'_2 = \{a'_2 A_2 \rightarrow a'_2 A_2 + a_1 A_1 \mid a_1, a'_2 > 0\}$ ,
- les réactions produisant  $A_1$ , sans  $A_2$  :  $R_1 \setminus R'_2 = \{s \rightarrow s + a_1 A_1 \mid A_1, A_2 \notin s, a_1 > 0\}$ ,
- les réactions produisant  $A_2$ , avec  $A_1$  en réactant :  $R_2 \cap R'_1 = \{a'_1 A_1 \rightarrow a'_1 A_1 + a_2 A_2 \mid a_2, a'_1 > 0\}$ ,
- les réactions produisant  $A_2$ , sans  $A_1$  :  $R_2 \setminus R'_1 = \{s \rightarrow s + a_2 A_2 \mid A_1, A_2 \notin s, a_2 > 0\}$ ,
- les réactions avec  $A_1$  en réactant et sans  $A_2$  :  $R'_1 \setminus R_2 = \{a'_1 A_1 \rightarrow s' \mid A_2 \notin s, a'_1 > 0\}$ ,
- les réactions avec  $A_2$  en réactant et sans  $A_1$  :  $R'_2 \setminus R_1 = \{a'_2 A_2 \rightarrow s' \mid A_1 \notin s, a'_2 > 0\}$ .

En appliquant  $D_1$ , les réactions  $r_2$ ,  $R_2 \setminus R'_1$  et  $R'_2 \setminus R_1$ , qui ne contiennent pas  $A_1$ , ne sont pas modifiées. La réaction  $r_1$  est supprimée. Les réactions de  $R_1$  sont combinées avec celles de  $R'_1$ . On obtient donc les réactions suivantes :

- (a)  $r_2 = A_2 \rightarrow \emptyset$ ,
- (b)  $R_2 \setminus R'_1 = \{s \rightarrow s + a_2 A_2 \mid A_1, A_2 \notin s, a_2 > 0\}$ ,
- (c)  $R'_2 \setminus R_1 = \{a'_2 A_2 \rightarrow s' \mid A_1 \notin s, a'_2 > 0\}$ ,
- (d)  $(R_1 \cap R'_2) \diamond_{A_1} (R_2 \cap R'_1)$ . Ces réactions sont de forme  $a'_2 A_2 \rightarrow (a_2 + a'_2) A_2$ ,
- (e)  $(R_1 \cap R'_2) \diamond_{A_1} (R'_1 \setminus R_2)$ . Ces réactions sont de forme  $a'_2 A_2 \rightarrow a'_2 A_2 + s'$ , avec  $A_2 \notin s'$ ,
- (f)  $(R_1 \setminus R'_2) \diamond_{A_1} (R_2 \cap R'_1)$ . Ces réactions sont de forme  $s \rightarrow s + a_2 A_2$ , avec  $A_2 \notin s$ ,
- (g)  $(R_1 \setminus R'_2) \diamond_{A_1} (R'_1 \setminus R_2)$ . Ces réactions, de forme  $s \rightarrow s + s'$ , ne contiennent pas de molécules de  $A_2$ .

On peut alors toujours utiliser  $D_2$ , puisque les ensembles (b) et (f) produisent des molécules de  $A_2$ , (a) les dégradent, et les réactions de (c), (d) et (e) n'ont que des molécules de  $A_2$  en réactant. On obtient les ensembles de réactions suivants :

- $(R_1 \setminus R'_2) \diamond_{A_1} (R'_1 \setminus R_2)$ ,
- $(R_2 \setminus R'_1) \diamond_{A_2} (R'_2 \setminus R_1)$ ,
- $(R_2 \setminus R'_1) \diamond_{A_2} ((R_1 \cap R'_2) \diamond_{A_1} (R'_1 \setminus R_2))$ ,
- $((R_1 \setminus R'_2) \diamond_{A_1} (R_2 \cap R'_1)) \diamond_{A_2} (R'_2 \setminus R_1)$ ,
- $(R_2 \setminus R'_1) \diamond_{A_2} ((R_1 \cap R'_2) \diamond_{A_1} (R_2 \cap R'_1))$ . Ces réactions sont toutes de forme  $s \rightarrow s$ , et peuvent être supprimées par la règle (RÉACTION CYCLIQUE),
- $((R_1 \setminus R'_2) \diamond_{A_1} (R_2 \cap R'_1)) \diamond_{A_2} ((R_1 \cap R'_2) \diamond_{A_1} (R_2 \cap R'_1))$ . Ces règles peuvent aussi être supprimées par (RÉACTION CYCLIQUE),
- $((R_1 \setminus R'_2) \diamond_{A_1} (R_2 \cap R'_1)) \diamond_{A_2} ((R_1 \cap R'_2) \diamond_{A_1} (R'_1 \setminus R_2))$ . Pour toute règle  $r = s \rightarrow s + a_1 A_1 \in R_1 \setminus R'_2$ , et  $r' = a' A_1 \rightarrow s' \in R'_1 \setminus R_2$ , la réaction correspondante dans cette ensemble sera  $s \rightarrow s + s' \setminus A_1$ . Mais cette réaction est déjà présente dans l'ensemble  $(R_1 \setminus R'_2) \diamond_{A_1} (R'_1 \setminus R_2)$ . Donc on peut supprimer cet ensemble.

Ainsi, on obtient uniquement les 4 premiers ensembles ci-dessus. Comme  $A_1 \notin R_2 \setminus R'_1$  et  $A_2 \notin R'_1 \setminus R_2$ , on a :

$$(R_2 \setminus R'_1) \diamond_{A_2} ((R_1 \cap R'_2) \diamond_{A_1} (R'_1 \setminus R_2)) = ((R_2 \setminus R'_1) \diamond_{A_2} (R_1 \cap R'_2)) \diamond_{A_1} (R'_1 \setminus R_2)$$

Donc, par symétrie, on obtient exactement les même réactions en commençant par  $D_2$ , suivie de  $D_1$  et (RÉACTION CYCLIQUE).

Donc cet ensemble de réactions est bien confluent. □

En revanche, notons que la simplification n'est ni fortement ni uniformément confluente. Considérons par exemple le réseau  $N = \{B \rightarrow B + A, C \rightarrow C + A, A \rightarrow A, A \rightarrow \emptyset\} \mid \emptyset$ , avec une fonction d'observation telle que  $\Omega(A) = \mathbb{O}$ , et  $A \in \mathcal{I}$ . On peut utiliser la règle (**RÉACTION CYCLIQUE**) et obtenir le réseau  $N_1 = \{B \rightarrow B + A, C \rightarrow C + A, A \rightarrow \emptyset\} \mid \emptyset$ , ou la règle (**CASCADE**) et obtenir le réseau  $N_2 = \{B \rightarrow B, C \rightarrow C\} \mid \emptyset$ . Ces deux réseaux peuvent être simplifiés pour obtenir le même réseau  $N_3 = \emptyset \mid \emptyset$ , mais il faut pour cela utiliser deux fois la règle (**RÉACTION CYCLIQUE**) sur  $N_2$  (et une fois (**CASCADE**) sur  $N_1$ ). La confluence n'est donc pas uniforme, ni forte.

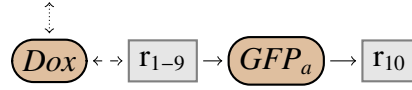
**Complexité** Enfin, on va brièvement étudier la complexité de nos différentes règles de simplification. On note  $n$  le nombre de réactions du réseau initial, et  $m$  le nombre d'espèces internes de celui-ci. La règle (**RÉACTION CYCLIQUE**) doit juste parcourir une fois l'ensemble des réactions, et a donc une complexité linéaire en  $n$ . Notons qu'en pratique, on élimine en un seul passage toutes les réactions cycliques, et pas une seule à la fois. Pour la règle (**ACTIVATEUR PRÉSENT**), appliquée sur une molécule spécifique  $A$ , on doit parcourir l'ensemble des réactions une première fois pour vérifier que le nombre de molécules de  $A$  ne décroît jamais, puis on simplifie une réaction. Encore une fois, on peut simplifier plusieurs réactions en un seul passage, afin de diminuer le temps de calcul. Cette règle est donc linéaire en  $n$ . C'est également le cas pour les règles (**ESPÈCE INUTILE**) et (**RÉACTION INVERSIBLE**). Enfin, considérons la règle (**CASCADE**), appliqué sur une molécule  $A$ . Comme on doit combiner des paires de réactions, on va cette fois être quadratique en  $n$ . La complexité sera donc de l'ordre de  $n^2$ . Comme la complexité donnée pour ces quatre règles est celle lorsque l'on connaît l'espèce  $A$  à simplifier, la complexité générale pour savoir si l'une des règles est applicable (ou si le réseau est complètement simplifié) est donc en  $n^2m$ .

Notons que la règle (**CASCADE**) peut augmenter de façon quadratique le nombre de réactions du réseau à chaque application. Cela augmente donc la complexité d'application des autres règles. Puisque la simplification est confluente, une bonne stratégie est donc de toujours essayer d'appliquer d'abord les autres règles avant (**CASCADE**). Le pire des cas d'un point de vue de la complexité serait une succession de (**CASCADE**), avec une complexité maximale pour la simplification de l'ordre de  $n^{2^m}$ .

### 3.2.3 Application au réseau Tet-On

On va maintenant appliquer ces règles de simplification sur le système Tet-On, illustré dans l'Ex. 8. Rappelons que le réseau est défini par :

$$N_{\text{Tet-On}} = R \mid Dox + 45000rtTA + 5P_{TRE3G}$$

Fig 3.3: Graphe du réseau simplifié  $N'_{\text{Tet-On}}$ .

avec  $\mathcal{I} = \{Dox_i, rtTA, rtTADox, P_{TRE3G}, mRNA, GFP, GFP_a\}$  et  $\mathcal{R}$  est l'ensemble des réactions décrites dans la Fig. 2.8. On simplifie ce réseau en utilisant la fonction d'observation  $\Omega_{\text{Tet-On}}$  présentée dans l'Ex. 11, qui ne compte que les molécules de  $GFP_a$  :  $\Omega_{\text{Tet-On}}(iGFP_a + s) = iGFP_a$ , avec  $GFP_a \notin s$ .

On va montrer que l'on peut simplifier ce réseau en utilisant la relation  $\Rightarrow_{\Omega_{\text{Tet-On}}}$ , et obtenir le réseau équivalent  $N'_{\text{Tet-On}}$  illustré dans la Fig. 3.3, et défini par :

$$N'_{\text{Tet-On}} = \{Dox \rightarrow Dox + GFP_a, GFP_a \rightarrow \emptyset\} \mid Dox.$$

Considérons d'abord la réaction  $r_3 = Dox_i + rtTA \rightarrow rtTADox$  et la réaction  $r_4 = rtTADox \rightarrow Dox_i + rtTA$ . Elles forment une paire de réactions inversibles. On a  $\Omega(rtTADox) = \emptyset = \Omega(Dox_i + rtTA)$ . De plus,  $rtTADox$  est uniquement utilisée en tant qu'activateur dans les autres réactions (ici la réaction  $r_5$ ). On peut alors utiliser la règle (**RÉACTION INVERSIBLE**), supprimer ces deux réactions, et remplacer dans tout le réseau les molécules de  $rtTADox$  par la solution  $Dox_i + rtTA$ . On remplace donc la réaction  $r_5$  par :

$$r_{3-5} = P_{TRE3G} + Dox_i + rtTA \rightarrow P_{TRE3G} + Dox_i + rtTA + mRNA.$$

La réaction  $r_{3-5}$  est alors la seule réaction utilisant les espèces  $rtTA$  et  $P_{TRE3G}$ , en tant qu'activateurs. Comme ces deux espèces sont présentes en quantité suffisante dans la solution initiale pour activer cette réaction (respectivement 45000 molécules de  $rtTA$  et 5 copies du gène  $P_{TRE3G}$ ), et qu'elles ne peuvent être modifiées par le contexte, on peut utiliser deux fois la règle (**ACTIVATEUR PRÉSENT**) et les enlever de cette réaction. On obtient alors la réaction :

$$r_{3-5}' = Dox_i \rightarrow Dox_i + mRNA.$$

Comme  $rtTA$  et  $P_{TRE3G}$  ne sont plus utilisées dans aucune réaction, qu'elles sont des espèces internes et ne sont pas observables, on peut les supprimer en utilisant la règle (**ESPÈCE INUTILE**) sur chacune d'elles. On obtient le réseau intermédiaire illustré dans la Fig. 3.4.

Ensuite, l'espèce  $mRNA$  est non observable et interne, et n'interagit que dans 3 réactions : la réaction  $r_{3-5}'$  qui la produit avec l'activateur  $Dox_i$ , la réaction  $r_6$  qui la dégrade, et la réaction  $r_7$  qui l'utilise comme activateur pour produire une

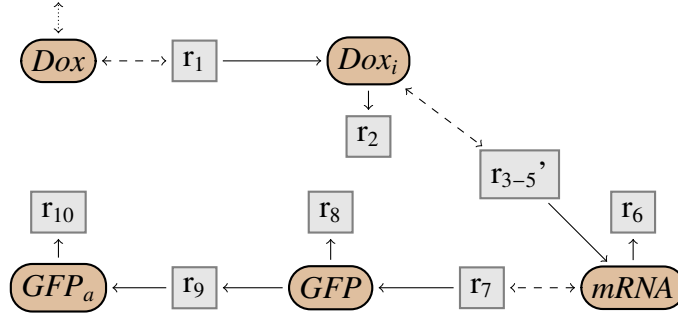


Fig 3.4: Graphe du réseau intermédiaire lors de la simplification de Tet-On.

molécule de  $GFP$ . On peut donc appliquer la règle (**CASCADE**), qui ici va supprimer l'espèce  $mRNA$  et fusionner ces trois réactions en une réaction qui produit directement  $GFP$  avec  $Dox_i$  comme activateur :

$$r_{3-7} = Dox_i \rightarrow Dox_i + GFP.$$

De manière similaire, on peut utiliser (**CASCADE**) sur l'espèce  $GFP$ , puis sur l'espèce  $Dox_i$ . On fusionne alors les réactions  $r_1$ ,  $r_2$ ,  $r_{3-7}$ ,  $r_8$  et  $r_9$  en la réaction :

$$r_{1-9} = Dox \rightarrow Dox + GFP_a.$$

On obtient alors le réseau  $N'_{Tet-On}$  illustré dans la Fig. 3.3.

D'après le Théorème 1, les deux réseaux sont bien attracteurs-congruents pour la fonction d'observation  $\Omega_{Tet-On}$ . Et notons que comme la relation de simplification est confluente d'après le Théorème 2, on aurait pu utiliser les règles dans un autre ordre, par exemple en commençant par appliquer (**ACTIVATEUR PRÉSENT**) sur  $P_{TRE3G}$  ou (**CASCADE**) sur  $GFP$ , et obtenir le même réseau complètement simplifié  $N'_{Tet-On}$ .

### 3.3 Règles de simplification plus générales

On a vu dans la section précédente un ensemble de règles de simplification, permettant notamment de simplifier le système Tet-On de façon confluente. Dans cette section, on enrichit la simplification avec d'autres règles. Certaines sont nouvelles, tandis que d'autres généralisent des règles de la section précédente. On montre que ces règles sont toutes correctes pour la congruence des attracteurs, et qu'elles diminuent la taille du réseau.

De façon générale, on peut classer les règles de simplification en quatre catégories. Une première catégorie regroupe les règles de simplification visant à supprimer une molécule "intermédiaire", c'est-à-dire qui n'est qu'une étape (non-observable) dans l'évolution du système. On regroupe ensuite des règles qui peuvent supprimer certaines réactions, lorsque celles-ci sont imitables en utilisant d'autres réactions du réseau. Une troisième catégorie contient des règles de simplification basées sur des symétries entre espèces dans le réseau. Enfin, la dernière catégorie regroupe tous les autres règles qui ne correspondent pas aux catégories précédentes.

#### 3.3.1 Règles d'élimination des espèces intermédiaires

On considère tout d'abord les règles visant à supprimer une espèce intermédiaire. Par exemple, considérons une espèce interne  $A$  qui ne serait utilisée uniquement dans les deux réactions  $r = s \rightarrow A$  et  $r' = A \rightarrow s'$ . On peut alors supprimer l'espèce  $A$ , et remplacer ces deux réactions par la réaction  $r \diamond_A r' = s \rightarrow s'$ .

On cherche ici à généraliser au maximum ce type de règles. La règle la plus générale combinerait toute réaction de production de forme  $s_1 \rightarrow s_2 + nA$  avec toute réaction de consommation de forme  $s'_1 + n'A \rightarrow s'_2$ . Néanmoins, cette simplification n'est pas correcte en générale, et nécessite d'ajouter des restrictions.

La Fig. 3.5 présente cinq règles de simplification d'espèces intermédiaires, décrites ci-dessous. L'exemple 29 explique alors pourquoi les restrictions sont nécessaires.

**Espèce intermédiaire présente que initialement** Dans la règle (**INTERMÉDIAIRE - INITIALE**), l'espèce  $A$  n'est présente que initialement, c'est-à-dire qu'aucune réaction ne peut créer de nouvelles molécules de  $A$ . De plus, il n'existe qu'une unique réaction pouvant consommer les molécules de  $A$ , en les transformant en une solution  $s$  : la réaction  $aA \rightarrow s$ . On peut alors directement remplacer les molécules de  $A$  par  $s$  (plus précisément, chaque ensemble de  $aA$  molécules est remplacé par une occurrence de  $s$ ), et supprimer cette réaction, sans modifier le comportement final du réseau. Si  $a' \neq 0$ , il peut alors rester des molécules de  $A$  dans la solution, qui

(INTERMÉDIAIRE - INITIALE)

$$\frac{A \in \mathcal{I} \quad A \notin s \quad a' < a}{\{aA \rightarrow s\} \mid (ka + a')A \Rightarrow \emptyset \mid ks + a'A}$$

(INTERMÉDIAIRE - SOLUTION)

$$\frac{A \in \mathcal{I} \quad \Omega(A) = \Omega(s) \quad A \notin s \quad \forall s_1 \rightarrow s_2 \in R. A \notin s_1}{R \cup \{A \rightarrow s\} \mid aA \Rightarrow R[A := s] \mid as}$$

(INTERMÉDIAIRE - RÉACTANT)

$$\frac{A \in \mathcal{I} \quad \forall s_1 \rightarrow s_2 \in R. A \notin s_1 \wedge A \in s_2 \quad R' \neq \emptyset \quad \forall r' \in R'. \exists s'. r' = A \rightarrow s' \wedge A \notin s' \wedge \Omega(s') = \Omega(A)}{R \cup R' \mid \emptyset \Rightarrow R \triangleright_A R' \mid \emptyset}$$

(INTERMÉDIAIRE - PRODUIT)

$$\frac{A \in \mathcal{I} \quad \forall r \in R. \exists s. r = s \rightarrow A \wedge A \notin s \wedge \Omega(s) = \Omega(A) \quad \forall s_1 \rightarrow s_2 \in R'. A \in s_1 \wedge A \notin s_2 \quad \exists s'. A \rightarrow s' \in R'}{R \cup R' \mid \emptyset \Rightarrow R \triangleleft_A R' \mid \emptyset}$$

(RÉDUCTION INVERSIBLE)

$$\frac{A \in \mathcal{I} \quad \Omega(A) = \Omega(s) \quad A \notin s \quad (R(N) \mid A) \rightarrow^* (R(N) \mid s) \quad (R(N) \mid s) \rightarrow^* (R(N) \mid A)}{N \Rightarrow N[A := s]}$$

Fig 3.5: Règles de simplification d'espèces intermédiaires pour la congruence des attracteurs, avec la fonction d'observation  $\Omega$  et les espèces internes  $\mathcal{I}$ .



peuvent éventuellement être supprimées par la règle (**ESPÈCE INUTILE**) si  $A$  n'est pas observable.

**Espèce intermédiaire présente dans la solution initiale** Avec la règle (**INTERMÉDIAIRE - SOLUTION**), le réseau a encore une fois des molécules de  $A$  dans sa solution, et une unique réaction  $A \rightarrow s$  pouvant les consommer. Mais cette fois, d'autres réactions peuvent créer des molécules de  $A$ . Sous la condition que  $A$  et  $s$  aient la même observation (comme expliqué dans l'exemple ci-dessous), on peut alors directement produire du  $s$  à la place des molécules de  $A$ . On remplace donc les réactions de  $R$  par  $R[A := s]$ . On remplace également les molécules présentes dans la solution initiale, et on supprime la réaction  $A \rightarrow s$ .

Notons que l'ensemble l'ensemble  $R$  pourrait ici être vide. Dans ce cas, cette règle serait similaire à la règle précédente, modulo la condition sur l'observation, avec  $a = 1$ . On pourrait également étendre cette règle en considérant plutôt une réaction de consommation de forme  $a'A \rightarrow s$ . Dans ce cas là, il faut imposer comme condition que toutes les réactions de  $R$  produisent un nombre de molécules de  $A$  multiple de  $a'$ , et de même que  $a$  soit multiple de  $a'$ . Mais dans ce cas, on peut alors appliquer la règle (**SOLUTION INDIVISIBLE**) (voir Fig. 3.11 plus bas), qui remplacerait  $aA$  par  $A$  partout, et donc ensuite appliquer la (**INTERMÉDIAIRE - SOLUTION**) dans sa forme actuelle.

**Espèce intermédiaire seule en réactant** Dans la règle (**INTERMÉDIAIRE - RÉACTANT**), chaque réaction consommant du  $A$  doit être de forme  $A \rightarrow s'$ , c'est-à-dire que ses réactants doivent être une unique molécule de  $A$ . On note  $R'$  l'ensemble de ces réactions. On va alors combiner chaque réaction  $r = s_1 \rightarrow s_2 + aA \in R$  produisant  $a$  molécules de  $A$  avec les réactions de  $R'$  de la façon suivante, afin d'obtenir toutes les combinaisons possibles :

$$r \triangleright_A R' = \{s_1 \rightarrow s_2 + s \mid s = \sum_{1 \leq i \leq a} s'_i, \exists r'_i = A \rightarrow s'_i \in R'\}$$

Par exemple, si  $r = \emptyset \rightarrow 2A$  et  $R' = \{A \rightarrow B, A \rightarrow C\}$ , on aura l'ensemble  $r \triangleright_A R' = \{\emptyset \rightarrow 2B, \emptyset \rightarrow B + C, \emptyset \rightarrow 2C\}$ . On étend cet opérateur à l'ensemble  $R$  des réactions produisant du  $A$ , et on remplace alors les réactions  $R \cup R'$  par  $R \triangleright_A R'$ .

**Espèce intermédiaire seule en produit** Le cas symétrique est traité par la règle (**INTERMÉDIAIRE - PRODUIT**). Toutes les réactions de production  $R$  sont de forme  $s \rightarrow A$ , c'est-à-dire avec une unique molécule de  $A$  comme produit. Pour toute

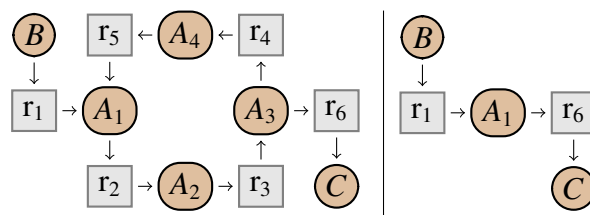


Fig 3.6: Simplification de cycles avec les règles (**RÉDUCTION INVERSIBLE**) et (**RÉACTION CYCLIQUE**).

réaction  $r' = s'_1 + a'A \rightarrow s'_2 \in R'$ , on définit l'ensemble des réactions suivant :

$$R \triangleleft_A R' = \{s'_1 + s \rightarrow s'_2 \mid s = \sum_{1 \leq i \leq a'} s_i\}.$$

$$\exists r_i = s_i \rightarrow A \in R$$

Comme précédemment, on remplace alors  $R \cup R'$  par  $R \triangleleft_A R'$ .

Dans ces deux règles, si l'ensemble  $R$  des réactions produisant du  $A$  est vide, alors  $R \triangleright_A R'$  et  $R \triangleleft_A R'$  sont également vides. Donc dans ce cas ces règles suppriment les réactions de  $R'$ , qui ne sont jamais utilisables, comme dans la règle (**RÉACTIONS INAPPLICABLES**) que nous verrons plus tard. L'ensemble  $R'$  quant à lui ne peut jamais être vide pour pouvoir appliquer l'une de ces deux règles.

**Réduction Inversible** Enfin, on peut généraliser la règle (**RÉACTION INVERSIBLE**) de la Fig. 3.2. En effet, certaines de ses conditions ne sont là que pour garantir la confluence, mais ne sont pas nécessaires pour qu'elle soit correcte. Par exemple, il n'est pas nécessaire que  $A$  ne soit utilisée ailleurs uniquement en tant qu'activateur. De même, il n'y a en fait pas besoin d'une paire de réactions inversibles directes entre  $A$  et  $s$ . Il faut juste que l'on puisse réduire  $A$  en  $s$  et réciproquement, c'est-à-dire  $(R(N) \mid A) \rightarrow^* (R(N) \mid s)$  et  $(R(N) \mid s) \rightarrow^* (R(N) \mid A)$ , comme le montre la règle généralisée (**RÉDUCTION INVERSIBLE**). Notons que cette règle n'est plus purement structurelle, puisqu'il est nécessaire de vérifier des conditions d'accessibilité pour pouvoir l'appliquer.

Cette règle permet notamment de simplifier des cycles (ou des composantes fortement connexes) ne contenant que des réactions mono-moléculaires (c'est-à-dire de forme  $A \rightarrow B$ ) et des espèces internes avec la même observation. Par exemple, les réactions présentées dans la Fig. 3.6 (gauche) peuvent être simplifiées avec cette règle pour obtenir les réactions de droite (après élimination des réactions cycliques), en supposant que les espèces  $A_i$  soient internes et aient toutes la même observation.

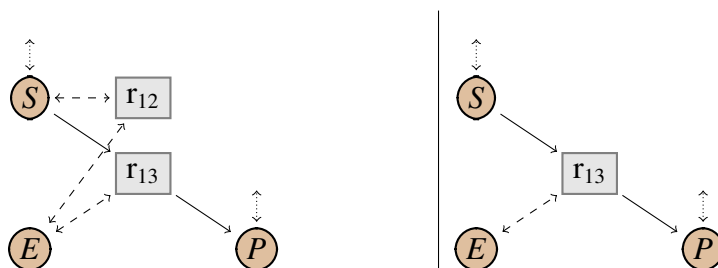


Fig 3.7: Graphes des réseaux intermédiaires lors de la simplification de Michaelis-Menten.

Contrairement à la règle (**RÉACTION INVERSIBLE**), la règle généralisée n'est plus confluente. Considérons le réseau  $\{A \rightarrow s, s \rightarrow A, A \rightarrow s', s' \rightarrow A, A \rightarrow B\} \mid \emptyset$ . On peut alors remplacer  $A$  soit par  $s$ , et obtenir le réseau  $\{s \rightarrow s', s' \rightarrow s, s \rightarrow B\} \mid \emptyset$ , soit par  $s'$ , et obtenir le réseau  $\{s \rightarrow s', s' \rightarrow s, s' \rightarrow B\} \mid \emptyset$ . Comme  $s$  et  $s'$  sont des solutions, interagissant potentiellement avec le contexte, on ne va pas pouvoir remplacer  $s$  par  $s'$  ou inversement, et donc on ne peut pas simplifier ces deux réseaux en un même réseau.

**Cascade** Enfin, notons que la règle (**CASCADE**) (Fig. 3.2) présentée dans la section précédente est aussi une règle d'élimination d'espèces intermédiaires, mais ne peut pas être généralisée plus.

**Exemple 28.** Considérons le réseau de Michaelis-Menten présenté dans l'Ex. 7 et décrit dans la Fig. 2.5. On a  $(R \mid E + S) \rightarrow^* (R \mid C)$  et  $(R \mid C) \rightarrow^* (R \mid E + S)$ . De plus  $C$  et  $E + S$  on la même observation, et  $C$  est une espèce interne. On peut appliquer la règle (**RÉDUCTION INVERSIBLE**), remplacer  $C$  par  $E + S$ , et obtenir le réseau simplifié illustré à gauche dans la Fig. 3.7 (notons qu'on a fusionné les deux réactions simplifiées identiques  $r_1$  et  $r_2$  en une réaction  $r_{12}$ , puisqu'on considère des ensembles de réactions, et non des multi-ensembles). On peut alors utiliser la règle (**RÉACTION CYCLIQUE**) pour supprimer  $r_{12}$  et obtenir le réseau  $\{r_3\} \mid E + S$ , représenté à droite de la Fig. 3.7. L'espèce  $E$  a une molécule dans la solution initiales, est interne, et aucune réaction du réseau ne peut modifier sa quantité de molécules. Donc, avec la règle (**ACTIVATEUR PRÉSENT**), on peut supprimer  $E$  de la réaction  $r_{12}$ . L'espèce est alors inutile, et peut être enlever avec la règle (**ESPÈCE INUTILE**). On obtient alors le réseau  $N'_{MM} = \{S \rightarrow P\} \mid S$  représenté dans la Fig. 3.8.

Dans les exemples suivants, on illustre la nécessité des différentes conditions présentes dans les règles.

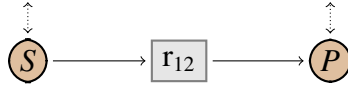
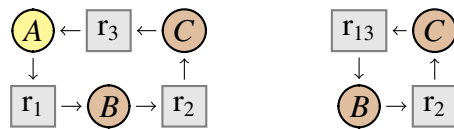


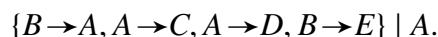
Fig 3.8: Graphe du réseau simplifié de Michaelis-Menten.

**Exemple 29.** Premièrement, remarquons que la plupart des règles contiennent une condition sur l'observation de  $A$ . En effet, dans ces règles, on cherche à remplacer une séquence de deux réactions (une produisant du  $A$  et l'autre le dégradant) en une seule réaction. Il faut alors que l'une de ces deux réactions ne soit pas observable pour obtenir les mêmes attracteurs. Par exemple, considérons le réseau  $\{B \rightarrow A, A \rightarrow C, C \rightarrow B\} \mid B$  (illustré graphiquement ci-dessous à gauche, avec en jaune l'espèce qu'on veut supprimer), que l'on veut simplifier en  $\{B \rightarrow C, C \rightarrow B\} \mid B$  (à droite) :

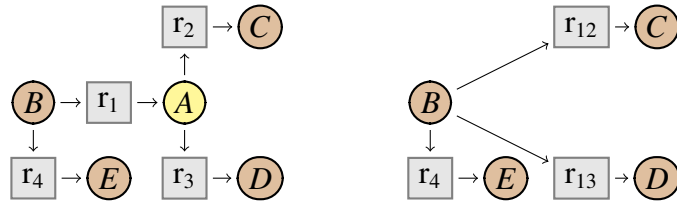


Le premier réseau peut converger vers  $\{\Omega(A), \Omega(B), \Omega(C)\}$ , tandis que le second peut converger vers  $\{\Omega(B), \Omega(C)\}$ . Pour être équivalent, il faut donc soit que  $\Omega(B) = \Omega(A)$ , soit  $\Omega(A) = \Omega(C)$ . Notons que dans la règle (**INTERMÉDIAIRE - INITIALE**), la condition d'observation n'est pas nécessaire, puisque l'espèce  $A$  ne sera présente que initialement (plus précisément, si un attracteur  $\mathcal{A}$  est accessible, et qu'il contient un réseau avec des molécules de  $A$ , alors par les propriétés de l'attracteur il contient également un réseau sans molécule de  $A$  (par fermeture) mais avec la même observation (par  $\Omega$ -connectivité). Il n'y a donc pas besoin d'ajouter de conditions supplémentaires.

Une autre restriction concerne la présence ou non de molécules de  $A$  dans la solution initiale du réseau. Si le réseau possède des molécules de  $A$  (règles (**INTERMÉDIAIRE - INITIALE**) et (**INTERMÉDIAIRE - SOLUTION**)), alors il doit n'y avoir qu'une unique réaction consommant du  $A$ . En effet, considérons le réseau suivant, représenté graphiquement à gauche ci-dessous :

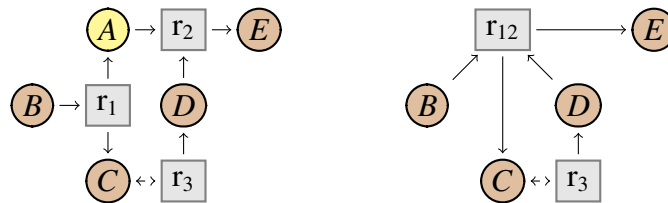
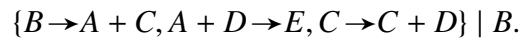


Sans la restriction ci-dessus, on voudrait pouvoir simplifier ce réseau, et obtenir les réactions  $\{B \rightarrow C, B \rightarrow D, B \rightarrow E\}$  (à droite).



Mais il faudrait également simplifier la molécule de  $A$  dans la solution initiale. On ne peut pas la transformer en une molécule de  $C$ , puisqu'on ne pourrait plus créer de molécules de  $D$  dans le réseau simplifié, alors que l'on pouvait le faire dans le réseau initial. Symétriquement, la molécule de  $A$  ne peut pas être transformée en  $D$ . Enfin, on ne peut pas non plus transformer  $A$  en  $B$ , sinon on est capable de produire du  $E$ , ce qu'on ne peut pas faire dans le réseau initial.

Remarquons aussi que dans toutes les règles (à part **RÉDUCTION INVERSIBLE** qui fonctionne différemment), soit toutes les règles de productions sont de forme  $s \rightarrow A$ , avec une unique molécule de  $A$  en produit, soit toutes les règles de consommation ont une unique molécule de  $A$  en réactant, donc sont de forme  $A \rightarrow s'$ . Sinon, considérons le réseau suivant (à gauche ci-dessous).



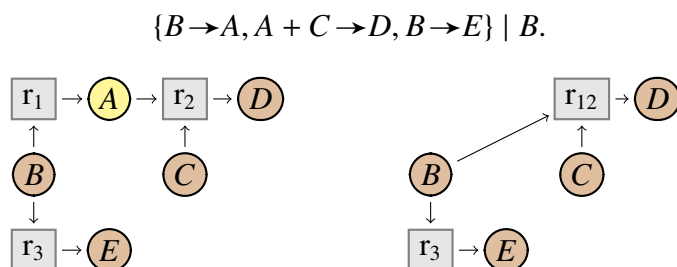
Dans ce réseau, on peut alors transformer  $B$  en  $A + C$ , puis produire une molécule de  $D$  avec l'activateur  $C$ , et transformer alors  $A + C + D$  en  $C + E$ . Le réseau simplifié serait alors (à droite ci-dessus) :



On ne peut donc appliquer aucune réaction, donc les deux réseaux ne sont pas équivalents. Le problème venant en fait du fait qu'on ne pouvait pas dans le réseau initial directement appliquer une réaction de production suivie d'une réaction de consommation, il était nécessaire d'appliquer une autre réaction entre les deux, ce qu'on ne peut pas faire dans le réseau simplifié.

La règle (**INTERMÉDIAIRE - PRODUIT**) est quasiment la symétrique de la règle (**INTERMÉDIAIRE - RÉACTANT**), mais possède une condition supplémentaire : la nécessité qu'il existe au moins une réaction de consommation de forme  $A \rightarrow s'$  (dans **INTERMÉDIAIRE - RÉACTANT**), cette condition découle directement des autres). Elle est

nécessaire pour ne pas se retrouver "bloqué" avec des molécules de  $A$ , qui ne pourraient être consommées. Considérons par exemple le réseau suivant, représenté graphiquement à gauche ci-dessous :

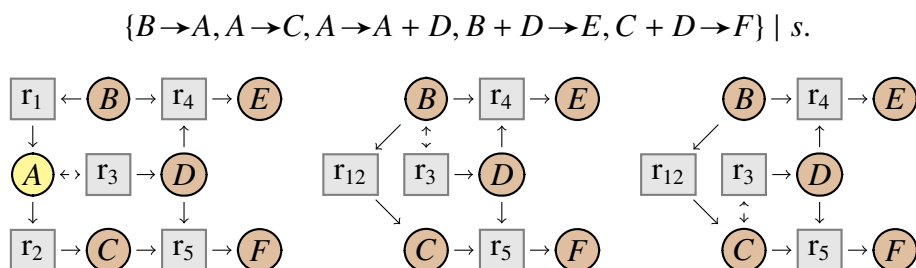


Il pourrait se simplifier pour obtenir le réseau de droite, défini par :



Dans le réseau initial, on peut transformer  $B$  en  $A$ , et se retrouver bloqué, sans pouvoir appliquer d'autres réactions. puisqu'on a pas de molécules de  $C$ . On ne peut pas non plus appliquer  $B \rightarrow E$ , puisque l'on a déjà transformé  $B$  en  $A$ . Donc on peut converger vers  $\{\Omega(A)\}$ . Dans le réseau simplifié par contre, on peut toujours appliquer cette réaction, et donc on doit toujours converger vers  $\{\Omega(E)\}$ . Donc, pour être équivalent, il faut toujours au moins une règle de consommation n'ayant que des molécules de  $A$  en réactant. Le même exemple en remplaçant la réaction  $A + C \rightarrow D$  par  $2A \rightarrow D$  montre qu'il ne doit y avoir qu'une seule molécule de  $A$  en réactant. En d'autres termes, il faut toujours une réaction de forme  $A \rightarrow s$ , pour une certaine solution  $s$ .

Enfin, remarquons que dans les règles (**INTERMÉDIAIRE - RÉACTANT**) et (**INTERMÉDIAIRE - PRODUIT**), l'espèce  $A$  ne peut jamais être utilisée comme activateur. Par exemple, considérons le réseau suivant, à gauche ci-dessous :



La combinaison des deux premières réactions produit la réaction  $B \rightarrow C$ . On peut alors remplacer la troisième réaction soit par  $B \rightarrow B + D$  (réseau du milieu ci-dessus), soit par  $C \rightarrow C + D$  (réseau de droite). Mais dans le premier cas, si on suppose  $s = B$ , on peut produire du  $E$  dans le réseau simplifié, mais pas dans le

$$\begin{array}{l}
 \text{(DÉCOMPOSITION - SUPPRESSION)} \quad \frac{(\mathbf{R} \mid s_1) \rightarrow^* (\mathbf{R} \mid s_2)}{\mathbf{R} \cup \{s_1 \rightarrow s_2\} \mid \emptyset \equiv \mathbf{R} \mid \emptyset} \\
 \\
 \text{(DÉCOMPOSITION - SIMPLIFICATION)} \quad \frac{\begin{array}{l} |s'_1 + s'_2| < |s_1 + s_2| \\ (\mathbf{R} \mid s_1) \rightarrow^* (\mathbf{R} \mid s'_1) \quad (\mathbf{R} \mid s'_1) \rightarrow^* (\mathbf{R} \mid s_1) \\ (\mathbf{R} \mid s_2) \rightarrow^* (\mathbf{R} \mid s'_2) \quad (\mathbf{R} \mid s'_2) \rightarrow^* (\mathbf{R} \mid s_2) \end{array}}{\mathbf{R} \cup \{s_1 \rightarrow s_2\} \mid \emptyset \equiv \mathbf{R} \cup \{s'_1 \rightarrow s'_2\} \mid \emptyset}
 \end{array}$$

Fig 3.9: Règles de décomposition pour la congruence des attracteurs, avec la fonction d'observation  $\Omega$  et les espèces internes  $\mathcal{I}$ .

réseau initial. Dans le deuxième cas, avec  $s = C$ , on peut produire du  $F$  dans le réseau simplifié, et pas dans le réseau initial.

Il existe néanmoins deux cas où  $A$  peut être en activateur. Premièrement le cas où toutes les réactions de production de  $A$  sont de forme  $s \rightarrow s + A$ , permettant alors de simplement remplacer  $A$  par  $s$  dans les réactions où il est activateur. Ce cas est déjà traité par la règle (**CASCADE**). Et Deuxièmement le cas où les réactions de production peuvent être inversées, comme dans la règle (**RÉDUCTION INVERSIBLE**).

### 3.3.2 Règles de décomposition

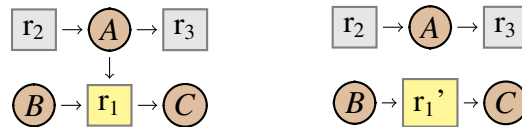
On considère maintenant deux règles de simplification qui visent respectivement à supprimer et à simplifier une réaction, lorsque celle-ci peut en fait être imitée par une combinaison d'autres réactions. Les deux règles sont présentées dans la Fig. 3.9. Notons que ces deux règles nécessitent de vérifier des conditions d'accessibilité, et ne sont donc pas purement structurelles.

**Suppression d'une réaction** Dans la première règle, (**DÉCOMPOSITION - SUPPRESSION**), on supprime complètement une réaction  $s_1 \rightarrow s_2$  dans le cas où on peut l'imiter avec un ensemble d'autres réactions. Par exemple, considérons le réseau  $\{A \rightarrow B, A \rightarrow C, C \rightarrow B\} \mid \emptyset$ , illustré graphiquement à gauche ci-dessous. Il peut être simplifié en  $\{A \rightarrow C, C \rightarrow B\} \mid \emptyset$  (à droite).



Notons que la règle (**RÉACTION CYCLIQUE**) est un cas particulier de (**DÉCOMPOSITION - SUPPRESSION**). Cette règle permet aussi par exemple de supprimer une réaction  $s_1 + s \rightarrow s_2 + s$  lorsqu'il y a déjà une réaction  $s_1 \rightarrow s_2$ .

**Simplification d'une réaction** La règle (**DÉCOMPOSITION - SIMPLIFICATION**) vise quant à elle à simplifier une réaction  $s_1 \rightarrow s_2$  en  $s'_1 \rightarrow s'_2$ , si on peut transformer  $s_i$  en  $s'_i$  et réciproquement à l'aide d'autres réactions. Par exemple, considérons le réseau  $\{A + B \rightarrow C, \emptyset \rightarrow A, A \rightarrow \emptyset\} \mid \emptyset$  et sa simplification  $\{B \rightarrow C, \emptyset \rightarrow A, A \rightarrow \emptyset\} \mid \emptyset$ .



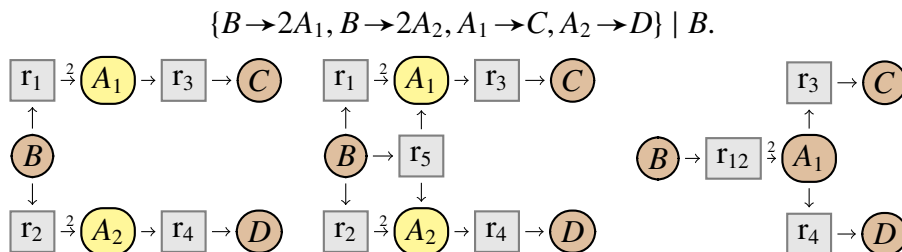
Dans le réseau initial, on peut imiter la réaction  $B \rightarrow C$  en créant une molécule de  $A$  puis en appliquant la réaction  $A + B \rightarrow C$ . Réciproquement dans le réseau simplifié, on peut imiter  $A + B \rightarrow C$  en appliquant  $B \rightarrow C$  et en supprimant la molécule de  $A$ .

### 3.3.3 Règles de symétrie

On considère maintenant le cas de règles basées sur les symétries entre deux espèces moléculaires. Elles doivent avoir la même observation, et un comportement "similaire". Dans le réseau  $\{B \rightarrow A_1, B \rightarrow A_2, A_1 \rightarrow C, A_2 \rightarrow C\} \mid \emptyset$  par exemple, les espèces  $A_1$  et  $A_2$  sont produites et consommées de façon identique. On peut donc les fusionner et obtenir le réseau  $\{B \rightarrow A_1, A_1 \rightarrow C\} \mid \emptyset$ .

On considère ici trois cas avec des conditions plus générales : soit les espèces symétriques sont toujours produites de façon identique, soit elles sont toujours consommées de façon identique, soit enfin tout comportement d'une espèce  $A$  peut être imité en remplaçant  $A$  par une solution moléculaire  $s$ . Les règles sont décrites formellement dans la Fig. 3.10. On verra dans la Sec. 3.4 qu'on peut par exemple utiliser ces règles pour simplifier différents réseaux du modèle MAPK.

Les espèces  $A_1$  et  $A_2$  peuvent être présentes dans une même solution moléculaire. Pour qu'elles soient symétriques, il faut alors considérer toutes les combinaisons possibles. Par exemple, considérons le réseau suivant, illustré à gauche ci-dessous :



Dans ce réseau, pour que  $A_1$  et  $A_2$  soient produites de façon symétrique, il faudrait ajouter la réaction  $B \rightarrow A_1 + A_2$  (réseau du milieu ci-dessus). Sinon, on peut atteindre la solution  $C + D$  dans le réseau simplifié (droite), mais pas dans le réseau de gauche.



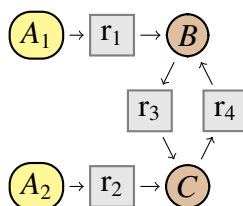
On définit formellement les notions de symétrie suivantes :

**Définition 13.** Soit  $A_1, A_2 \in \text{Espèces}$  deux espèces moléculaires,  $s, s' \in \text{Solutions}$  deux solutions, et  $R$  un ensemble de réactions. On dit que  $A_1$  et  $A_2$  sont :

- *symétriques dans*  $(s, s')$ , noté  $A_1 \overset{sol}{\underset{s, s'}{\rightleftarrows}} A_2$ , si le nombre total de  $A_1$  et  $A_2$  est le même dans les deux solutions, et que le reste de la solution est également identique :  $s(A_1) + s(A_2) = s'(A_1) + s'(A_2)$  et  $s \setminus \{A_1, A_2\} = s' \setminus \{A_1, A_2\}$ ,
- *réactants-symétriques dans*  $R$ , noté  $A_1 \overset{rea}{\underset{R}{\rightleftarrows}} A_2$ , si elles sont toujours consommées de façon identique, c'est-à-dire pour toute réaction  $s_1 \rightarrow s_2 \in R$ , pour toute solution  $s'_1$  telle que  $A_1 \overset{sol}{\underset{s_1, s'_1}{\rightleftarrows}} A_2$ , il existe une solution  $s'_2$  telle que  $A_1 \overset{sol}{\underset{s_2, s'_2}{\rightleftarrows}} A_2$  et que  $R \mid s'_1 \rightarrow^* R \mid s'_2$ ,
- *produits-symétriques dans*  $R$ , noté  $A_1 \overset{pro}{\underset{R}{\rightleftarrows}} A_2$ , si elles sont toujours produites de façon identique, c'est-à-dire pour toute réaction  $s_1 \rightarrow s_2 \in R$ , pour toute solution  $s'_2$  telle que  $A_1 \overset{sol}{\underset{s_2, s'_2}{\rightleftarrows}} A_2$ , il existe une solution  $s'_1$  telle que  $A_1 \overset{sol}{\underset{s_1, s'_1}{\rightleftarrows}} A_2$  et que  $R \mid s'_1 \rightarrow^* R \mid s'_2$ ,
- $A_1$  peut être *imitée* par  $s$  dans  $R$ , noté  $A_1 \overset{app}{\underset{R}{\rightleftarrows}} s$ , si pour toute réaction  $s_1 \rightarrow s_2 \in R$ , on peut faire  $R \mid s_1[A_1 := s] \rightarrow^* R \mid s_2[A_1 := s]$ .

Par exemple, dans les réseaux ci-dessus,  $A_1$  et  $A_2$  sont symétriques dans les solutions  $(2A_1, 2A_2)$ , ou  $(2A_1, A_1 + A_2)$ . Elles ne sont pas réactants-symétriques, puisque les réactions  $A_1 \rightarrow C$  et  $A_2 \rightarrow D$  produisent des solutions  $C$  et  $D$  qui ne sont pas symétriques. Elles sont produits-symétriques dans  $\{B \rightarrow 2A_1, B \rightarrow 2A_2, B \rightarrow A_1 + A_2\}$  (milieu), mais pas dans  $\{B \rightarrow 2A_1, B \rightarrow 2A_2\}$  (gauche).

On peut noter qu'une réaction  $A_1 \rightarrow B$  consommant (ou produisant) du  $A_1$  peut être imitée à l'aide d'une chaîne de réactions  $R \mid A_2 \rightarrow^* R \mid B$ . Ainsi,  $A_1$  et  $A_2$  sont réactant-symétriques dans  $\{A_1 \rightarrow B, A_2 \rightarrow C, B \rightarrow C, C \rightarrow B\}$ .

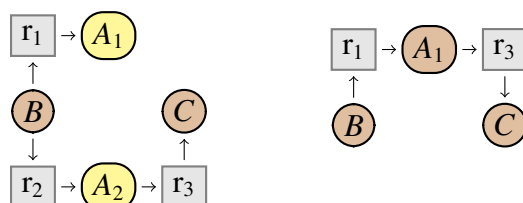


$$\begin{array}{l}
\text{(SYMÉTRIE - RÉACTANTS)} \\
\frac{A_1, A_2 \in \mathcal{I} \quad \Omega(A_1) = \Omega(A_2)}{A_1 \underset{\text{R}}{\overset{\text{rea}}{\rightleftharpoons}} A_2 \quad \forall s_1 \rightarrow s_2 \in \text{R}' . A_1 \underset{s_1, s_2}{\overset{\text{sol}}{\rightleftharpoons}} A_2} \\
\text{R} \cup \text{R}' \mid s \Rightarrow \text{R}[A_2 := A_1] \mid s[A_2 := A_1] \\
\\
\text{(SYMÉTRIE - PRODUITS)} \\
\frac{A_1, A_2 \in \mathcal{I} \quad \Omega(A_1) = \Omega(A_2)}{A_1 \underset{\text{R}}{\overset{\text{pro}}{\rightleftharpoons}} A_2 \quad \forall s_1 \rightarrow s_2 \in \text{R}' . A_1 \underset{s_1, s_2}{\overset{\text{sol}}{\rightleftharpoons}} A_2} \\
\exists s, s' . A_1 \rightarrow s \in \text{R} \wedge A_2 \rightarrow s' \in \text{R} \\
\text{R} \cup \text{R}' \mid \emptyset \Rightarrow \text{R}[A_2 := A_1] \mid \emptyset \\
\\
\text{(SYMÉTRIE - IMITATION)} \\
\frac{A \in \mathcal{I} \quad \Omega(A) = \Omega(s) \quad A \notin s}{A \underset{\text{R}}{\overset{\text{app}}{\rightleftharpoons}} s \quad \exists s' . A \rightarrow s' \in \text{R}} \\
\text{R} \mid \emptyset \Rightarrow \text{R}[A := s] \mid \emptyset
\end{array}$$

Fig 3.10: Règles de symétrie pour la congruence des attracteurs, avec la fonction d'observation  $\Omega$  et les espèces internes  $\mathcal{I}$ .

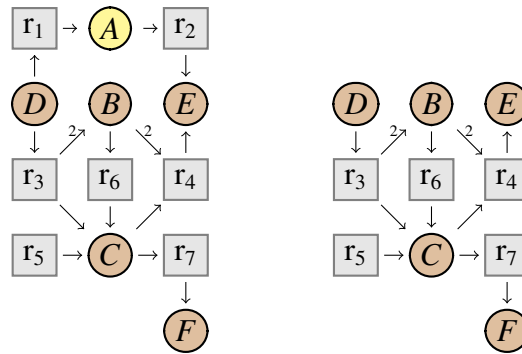
**Symétrie - Réactants** Dans la première règle, (**SYMÉTRIE - RÉACTANTS**), les deux espèces  $A_1$  et  $A_2$  sont réactants-symétriques dans  $\text{R}$ , et donc sont toujours consommées de façon identiques. L'ensemble  $\text{R}'$  de réactions transforme des molécules de  $A_1$  en  $A_2$ , sans changer leur quantité totale ni les autres molécules. On remplace alors toutes les occurrences de  $A_2$  par  $A_1$ . L'ensemble  $\text{R}$  peut être vide. Dans ce cas, les espèces  $A_1$  et  $A_2$  sont inutiles, puisque les seules réactions les utilisant sont celles de  $\text{R}'$ , qui transforment  $A_1$  en  $A_2$  ou réciproquement.

**Symétrie - Produits** La deuxième règle (**SYMÉTRIE - PRODUITS**) est similaire, pour le cas où  $A_1$  et  $A_2$  sont produits-symétriques. On doit dans ce cas ajouter deux restrictions supplémentaires. Premièrement,  $A_1$  et  $A_2$  ne doivent pas apparaître dans la solution du réseau initial. De plus, comme pour la règle (**INTERMÉDIAIRE - PRODUIT**), pour éviter que le réseau soit "bloqué" sur  $A_1$  ou  $A_2$ , il faut qu'il possède au moins une réaction de forme  $A_1 \rightarrow s$  et une réaction de forme  $A_2 \rightarrow s_2$ . Sinon, considérons le réseau ci-dessous (gauche), qui se simplifierait pour former le réseau de droite.



Dans le réseau de gauche, on peut transformer  $B$  en  $A_1$  puis être bloqué, tandis que ce n'est jamais possible dans le réseau de droite.

**Symétrie - Imitation** Enfin, dans la règle (**SYMÉTRIE - IMITATION**), tout comportement de l'espèce  $A$  peut être imité par une solution  $s$  de même observation, avec  $A \notin s$ . La solution  $s$  peut par contre avoir d'autres comportements. L'espèce  $A$  doit être interne et ne doit pas être dans la solution initiale du réseau. La condition  $\exists s'. A \rightarrow s' \in R$  implique que s'il y a du  $A$  dans la solution courante du réseau, alors il peut toujours se réduire avec la sémantique opérationnelle, il n'est jamais "bloqué" sur  $A$ . On peut par exemple simplifier le réseau de gauche ci-dessous et obtenir le réseau de droite, avec  $s = 2B+C$  (après fusion des réactions identiques) :



### 3.3.4 Autres règles

Enfin, la dernière catégorie regroupe les autres règles de simplification. Ces règles sont en général assez simples, modifiant que peu ou pas du tout le graphe des solutions accessibles des réseaux, comme les règles (**ACTIVATEUR PRÉSENT**) et (**ESPÈCE INUTILE**) (Fig. 3.2). Les nouvelles règles sont présentées dans la Fig. 3.11.

**Réactions inapplicables** Dans la règle (**RÉACTIONS INAPPLICABLES**), on considère une espèce  $A$ , interne, dont la quantité ne peut pas augmenter. On va alors supprimer l'ensemble des réactions d'un ensemble  $R_1$ , qui ne sont pas jamais applicables, car il n'y a pas assez de molécules de  $A$  ( $\forall s_1 \rightarrow s_2 \in R_1. s_1(A) > a_0$ ). Les réactions de l'ensemble  $R_2$  sont applicables, mais ne peuvent jamais augmenter le nombre de molécules de  $A$  (c'est-à-dire  $\forall r \in R_2. stoic_r(A) \leq 0$ ).

**Productions Groupées** Dans cette règle, l'ensemble  $R$  ne contient que des réactions de production de solutions moléculaires non observables, c'est-à-dire de forme  $\emptyset \rightarrow s_i$  avec  $\Omega(s_i) = \emptyset$ . On veut alors fusionner ces réactions en une unique réaction  $\emptyset \rightarrow s$ , avec  $s$  contenant une molécule de chaque espèce apparaissant dans

(RÉACTIONS INAPPLICABLES)

$$\frac{A \in \mathcal{I} \quad \mathbf{R}_1 \neq \emptyset \quad \forall s_1 \rightarrow s_2 \in \mathbf{R}_1. s_1(A) > a_0 \quad \forall \mathbf{r} \in \mathbf{R}_2. stoic_{\mathbf{r}}(A) \leq 0}{\mathbf{R}_1 \cup \mathbf{R}_2 \mid a_0 A \Rightarrow \mathbf{R}_2 \mid a_0 A}$$

(PRODUCTIONS GROUPÉES)

$$\frac{s = \sum_{A \in \text{Espèces}(\mathbf{R})} A \quad \forall \mathbf{r} \in \mathbf{R}. \exists s'. \mathbf{r} = \emptyset \rightarrow s' \wedge \Omega(s') = \mathbf{0} \quad |\mathbf{R}| > 1 \vee \exists \mathbf{r} \in \mathbf{R}. \exists A. stoic_{\mathbf{r}}(A) > 1}{\mathbf{R} \mid \emptyset \Rightarrow \{\emptyset \rightarrow s\} \mid \emptyset}$$

(SOLUTION INDIVISIBLE)

$$\frac{\text{Espèces}(s) \subseteq \mathcal{I} \quad A \in s \quad \Omega(s) = \Omega(A) \quad \text{Indiv}(s, s_0) \quad \text{Indiv}(s, \mathbf{R}) \quad |s| > 1}{\mathbf{R} \mid s_0 \Rightarrow \mathbf{R}[s := A] \mid s_0[s := A]}$$

(ESPÈCE DÉPENDANTE)

$$\frac{A, B \in \mathcal{I} \quad \Omega(A) = \mathbf{0} \quad a \geq b \quad \forall \mathbf{r} = s_1 \rightarrow s_2 \in \mathbf{R}. s_1(B) \geq s_1(A) \wedge stoic_{\mathbf{r}}(A) \geq stoic_{\mathbf{r}}(B)}{\mathbf{R} \mid aA + bB \Rightarrow \mathbf{R} \setminus A \mid bB}$$

Fig 3.11: Autres règles de simplification pour la congruence des attracteurs, avec la fonction d'observation  $\Omega$  et les espèces internes  $\mathcal{I}$ .

R. La condition  $|\mathcal{R}| > 1 \vee \exists r \in \mathcal{R}. \exists A. stoic_r(A) > 1$  garantit que la taille du réseau diminue : soit on fusionne plusieurs réactions, soit il y a une unique réaction, mais produisant plusieurs molécules d'une même espèce (par exemple, la règle peut simplifier une réaction  $\emptyset \rightarrow 2A + B$  en  $\emptyset \rightarrow A + B$ ).

**Solution Indivisible** Dans la règle (**SOLUTION INDIVISIBLE**), on considère une solution  $s$  *indivisible*, c'est-à-dire telle que toutes les espèces utilisées dans  $s$  sont toujours utilisées ensemble, et sont internes. Formellement, on dit que  $s$  est indivisible dans  $s'$ , noté  $Indiv(s, s')$ , s'il existe un entier  $a \geq 0$  et une solution  $s''$  telle que  $s' = as + s''$  avec  $s \cap s'' = \emptyset$ . On veut alors remplacer  $s$  par une de ses espèces  $A$ , avec la même observation :  $\Omega(s) = \Omega(A)$ . Pour toute solution  $s' = as + s''$  avec  $s \cap s'' = \emptyset$ , on définit  $s'[s := A] = aA + s''$ . De même, on dit que  $s$  est indivisible dans  $r = s_1 \rightarrow s_2$ , noté  $Indiv(s, r)$ , si  $Indiv(s, s_1)$  et  $Indiv(s, s_2)$ , et  $r[s := A] = s_1[s := A] \rightarrow s_2[s := A]$ . On étend ces définitions aux ensembles de réactions de façon naturelle. La dernière condition garantit que la taille du réseau diminue.

**Espèce Dépendante** Dans la règle (**ESPÈCE DÉPENDANTE**), une espèce  $A$  est toujours utilisée en réactant en même temps qu'une espèce  $B$ . De plus, il y a dans la solution initiale au moins autant de molécules de  $A$  que de  $B$ , et on produit toujours au moins autant de  $A$  que de  $B$ . On peut donc supprimer l'espèce  $A$ . En effet, si  $A$  n'est pas présente en quantité suffisante pour appliquer une réaction, alors c'est également le cas de  $B$ .

### 3.3.5 Propriétés

On prouve ici que ces règles réduisent bien la taille des réseaux et qu'elles sont correctes pour la congruence des attracteurs. On discute également de la décidabilité de l'application de ces règles.

**Simplification** On commence par montrer que nos règles simplifient bien la taille des réseaux.

**Proposition 5.** Pour toute règle de simplification  $D$  présentée dans cette section, si on peut simplifier un réseau  $N$  en un réseau  $N'$  en utilisant la règle  $D$ , alors  $N' < N$ .

*Proof.* Toutes les règles éliminant une espèce intermédiaire de la Fig. 3.5 réduisent trivialement le nombre d'espèces moléculaires. C'est également le cas de toutes les règles d'espèces symétriques présentées dans la Fig. 3.10, et de la règle (**ESPÈCE DÉPENDANTE**).

La règle (**DÉCOMPOSITION - SUPPRESSION**) diminue le nombre de réactions. C'est également le cas de la règle (**RÉACTIONS INAPPLICABLES**), d'après la condition  $R_1 \neq \emptyset$ .

La règle (**DÉCOMPOSITION - SIMPLIFICATION**) diminue la taille combinée des solutions du réseau, d'après la condition  $|s'_1 + s'_2| < |s_1 + s_2|$ .

La règle (**PRODUCTIONS GROUPÉES**) diminue soit le nombre de réactions (si  $|R| > 1$ ) ou la taille combinée du réseau (si  $\exists r \in R. \exists A. stoic_r(A) > 1$ ).

Enfin, la règle (**SOLUTION INDIVISIBLE**) diminue soit le nombre d'espèces soit la taille combinée du réseau.  $\square$

**Correction** On vérifie maintenant que nos règles sont bien correctes pour la congruence. On prouve tout d'abord les lemmes suivants. L'idée est de définir une fonction  $f$  entre les réseaux accessibles avant la simplification, et ceux accessibles après, telle que pour tout réseau  $N_1$ , le comportement (et l'observation) de  $N_1$  soit le même que celui de  $f(N_1)$ , et réciproquement avec une fonction  $g$ . L'existence de ces fonctions implique l'équivalence des attracteurs. Notons qu'en pratique, on aura en général  $g = f^{-1}$ . De plus, on a souvent directement  $f(N) = N'$  et  $g(N') = N$  au lieu des conditions 1 et 2 dans le lemme ci-dessous.

**Lemme 4.** Soit  $\Omega$  une fonction d'observation et  $N, N'$  deux réseaux de réactions. S'il existe deux fonctions  $f : \{N \rightarrow^* \} \mapsto \{N' \rightarrow^* \}$  et  $g : \{N' \rightarrow^* \} \mapsto \{N \rightarrow^* \}$  vérifiant les points suivants, alors  $N \simeq_{\Omega} N'$  :

1.  $N \rightarrow^* g(N')$ ,
2.  $N' \rightarrow^* f(N)$ .
3. pour tout réseau  $N_1, N_2 \in \{N \rightarrow^* \}$ ,  $N_1 \rightarrow^* N_2$  implique  $f(N_1) \rightarrow^* f(N_2)$ ,
4. pour tout réseau  $N'_1, N'_2 \in \{N' \rightarrow^* \}$ ,  $N'_1 \rightarrow^* N'_2$  implique  $g(N'_1) \rightarrow^* g(N'_2)$ ,
5. pour tout réseau  $N_1 \in \{N \rightarrow^* \}$ ,  $N_1 \rightarrow^* g(f(N_1))$ ,
6. pour tout réseau  $N'_1 \in \{N' \rightarrow^* \}$ ,  $N'_1 \rightarrow^* f(g(N'_1))$ ,
7.  $\Omega \circ f = \Omega = \Omega \circ g$ ,

*Proof.* On va montrer que pour tout ensemble d'observables  $O \subseteq Observables$ , on a :

$$N \downarrow_O^{\Omega} \quad \text{si et seulement si} \quad N' \downarrow_O^{\Omega} .$$

Soit  $\mathcal{A}$  un  $\Omega$ -attracteur accessible depuis  $N$ , et  $N_1 \in \mathcal{A}$  un réseau appartenant à l'attracteur. On pose  $\mathcal{A}' = \{f(N_1) \rightarrow^* \}$ . On a  $N \rightarrow^* N_1$ , donc  $N' \rightarrow^* f(N) \rightarrow^* f(N_1)$  (conditions 2 et 3). Donc  $\mathcal{A}'$  est accessible depuis  $N'$ . On va montrer que c'est un  $\Omega$ -attracteur. Il est fermé par définition, donc il faut montrer qu'il est  $\Omega$ -connecté.

Soit  $N'_1 \in \mathcal{A}'$ , et  $o \in \Omega(\mathcal{A}')$ . Il existe un réseau  $N'_2 \in \mathcal{A}'$  tel que  $\Omega(N'_2) = o$ . On a  $f(N_1) \rightarrow^* N'_2$ , donc  $N_1 \rightarrow^* g(f(N_1)) \rightarrow^* g(N'_2)$  (conditions 4 et 5). Donc  $g(N'_2) \in \mathcal{A}$ , et  $\Omega(g(N'_2)) = \Omega(N'_2) = o \in \Omega(\mathcal{A})$  (condition 7).

De même,  $f(N_1) \rightarrow^* N'_1$ , donc  $N_1 \rightarrow^* g(f(N_1)) \rightarrow^* g(N'_1)$ , et  $g(N'_1) \in \mathcal{A}$ . Or  $\mathcal{A}$  est un attracteur, donc est  $\Omega$ -connecté, donc  $o \in \Omega(\{g(N'_1) \rightarrow^*\})$ . Donc  $\exists N_2 \in \{g(N'_1) \rightarrow^*\}$  avec  $o = \Omega(N_2)$ . Donc  $N'_1 \rightarrow^* f(g(N'_1)) \rightarrow^* f(N_2)$ , et  $\Omega(f(N_2)) = o \in \Omega(\{N'_1 \rightarrow^*\})$  (conditions 3 et 6). Donc pour tout  $N'_1 \in \mathcal{A}'$ ,  $\Omega(\mathcal{A}') \subseteq \Omega(\{N'_1 \rightarrow^*\})$ , donc  $\mathcal{A}'$  est *obs*-connecté et c'est un attracteur. Comme l'observation est préservée, on a directement  $\Omega(\mathcal{A}') = \Omega(\mathcal{A})$ , donc  $N \downarrow_o^\Omega$  implique  $N' \downarrow_o^\Omega$ . Par symétrie, la réciproque est vraie aussi, et donc :

$$N \downarrow_o^\Omega \quad \text{si et seulement si} \quad N' \downarrow_o^\Omega .$$

Pour la divergence, si  $N \uparrow^\Omega$ , alors il existe un réseau  $N_1 \in \{N \rightarrow^*\}$  tel que  $\forall O \subseteq \text{Observables}$ ,  $N_1 \not\downarrow_o^\Omega$ . Mais, avec le même raisonnement que ci-dessus, cela implique  $\forall O \subseteq \text{Observables}$ ,  $f(N_1) \not\downarrow_o^\Omega$ , or  $N' \rightarrow^* f(N) \rightarrow^* f(N_1)$ , donc  $N' \uparrow^\Omega$ . La réciproque est symétrique, donc :

$$N \uparrow^\Omega \quad \text{si et seulement si} \quad N' \uparrow^\Omega .$$

Donc les deux réseaux sont  $\Omega$ -attracteurs-équivalents. □

Notons qu'on aurait pu utiliser ce lemme pour la preuve du Théoreme 1, pour les règles (ACTIVATEUR PRÉSENT) (avec  $g = f^{-1}$ ) et (RÉACTION INVERSIBLE). On présente maintenant une variante de ce lemme, où  $g(N')$  est un ensemble de réseaux. La preuve est omise, elle est similaire à celle de la règle (ESPÈCE INUTILE) dans le Théoreme 1, dont les fonctions  $f$  et  $g$  vérifient exactement les propriétés du lemme ci-dessous.

**Lemme 5.** Soit  $\Omega$  une fonction d'observation et  $N, N'$  deux réseaux de réactions. S'il existe deux fonctions  $f : \{N \rightarrow^*\} \rightarrow \{N' \rightarrow^*\}$  et  $g : \{N' \rightarrow^*\} \rightarrow 2^{\{N \rightarrow^*\}}$  vérifiant les points suivants, alors  $N \simeq_\Omega N'$  :

1.  $N \in g(N')$ ,
2.  $N' = f(N)$ .
3. pour tout réseau  $N_1, N_2 \in \{N \rightarrow^*\}$ ,  $N_1 \rightarrow^* N_2$  implique  $f(N_1) \rightarrow^* f(N_2)$ ,
4. pour tout réseau  $N'_1, N'_2 \in \{N' \rightarrow^*\}$ ,  $N'_1 \rightarrow^* N'_2$  implique  $\forall N_1 \in g(N'_1). \exists N_2 \in g(N'_2). N_1 \rightarrow^* N_2$ ,
5. pour tout réseau  $N_1 \in \{N \rightarrow^*\}$ ,  $N_1 \in g(f(N_1))$ ,
6. pour tout réseau  $N'_1 \in \{N' \rightarrow^*\}$ ,  $f(g(N'_1)) = \{N'_1\}$ ,

7.  $\Omega \circ f = \Omega$ , et pour tout  $N_1 \in g(N'_1)$ , on a  $\Omega(N_1) = \Omega(N'_1)$ ,

**Théorème 3.** Pour toute règle de simplification  $D$ , si on peut simplifier un réseau  $N$  en un réseau  $N'$  en utilisant la règle  $D$ , alors  $N' \sim_{\Omega, \mathcal{I}} N$ .

*Proof.* Soit  $N$  le réseau initial,  $N'$  le réseau simplifié, et  $N''$  un contexte. Les réactions du contexte  $N''$  seront notées  $R''$ . Les preuves utilisent principalement le Lem. 4, le Lem. 5 ou des variantes, entre les réseaux  $N \mid N''$  et  $N' \mid N''$ . Les preuves des règles (**INTERMÉDIAIRE - PRODUIT**) et (**INTERMÉDIAIRE - RÉACTANT**) sont un peu plus compliquées, et nécessitent de réordonner certaines séquences de réactions, comme pour la règle (**CASCADE**) dans le Théorème 1.

(**INTERMÉDIAIRE - INITIALE**) : On montre tout d'abord que  $N$  est congruent au réseau  $h(N)$  obtenu en appliquant récursivement la réaction  $aA \rightarrow s$ , puis que  $h(N)$  est en fait congruent à  $N'$ .

Tout réseau  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  est de forme  $N_1 = \{aA \rightarrow s\} \cup R'' \mid (ka + a')A + s''$ , avec  $a' < a$  et  $A \notin s''$ . On définit alors la fonction  $h(N_1) = \{aA \rightarrow s\} \cup R'' \mid ks + a'A + s''$ . On a trivialement  $N_1 \rightarrow^* h(N_1)$  en appliquant la réaction  $A \rightarrow s$ . De plus, pour tout  $N_2 \in \{h(N_1) \rightarrow^*\}$ ,  $A \notin s(N_2)$ . Pour tout attracteur  $\mathcal{A}$  accessible depuis  $N_1$ ,  $h(\mathcal{A}) \subseteq \mathcal{A}$  est aussi un attracteur, accessible depuis  $h(N_1)$ , avec la même observation. De même  $N_1$  peut diverger si et seulement si  $h(N_1)$  peut diverger. Donc  $N \mid N''$  et  $h(N \mid N'')$  sont équivalents.

On montre maintenant que  $h(N \mid N'')$  et  $N' \mid N''$  sont aussi équivalents. Tout réseau  $N_1 \in h(\{(N \mid N'') \rightarrow^*\})$  est de forme  $N_1 = \{A \rightarrow s\} \cup R'' \mid a'A + s''$  avec  $A \notin s''$ , on définit alors  $f(N_1) = R'' \mid a'A + s''$  et  $g = f^{-1}$ . Alors trivialement  $N_1 \rightarrow^* N_2$  si et seulement si  $f(N_1) \rightarrow^* f(N_2)$ . Donc pour tout  $\mathcal{A}$  attracteur accessible depuis  $N_1$ ,  $f(\mathcal{A})$  est aussi un attracteur avec la même observation. Réciproquement, pour tout  $\mathcal{A}'$ ,  $g(\mathcal{A}')$  est un attracteur avec la même observation. Les preuves pour la divergence sont triviales. Donc  $N \sim_{\Omega, \mathcal{I}} h(N) \sim_{\Omega, \mathcal{I}} N'$ .

(**INTERMÉDIAIRE - SOLUTION**) : Cette preuve utilise le Lem. 4.

Tout réseau  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  est de forme  $N_1 = R \cup \{A \rightarrow s\} \cup R'' \mid a_1A + s''$ , avec  $A \notin s''$  et  $R$  un ensemble de réactions sans molécule de  $A$  en réactant. On définit la fonction  $f(N_1) = R[A := s] \cup R'' \mid a_1s + s''$ , qui remplace les molécules de  $A$  par la solution  $s$ .

Le réseau  $N'$  est alors égale à  $R[A := s] \mid a_1s$  d'après la règle de simplification. Comme les réactions ne sont pas modifiées par la sémantique opérationnelle  $\rightarrow$ , tout réseau  $N'_1 \in \{(N' \mid N'') \rightarrow^*\}$  est de forme  $N'_1 = R[A := s] \cup R'' \mid s''$ , avec  $A \notin s''$  (avec  $R$  toujours les réactions de  $N_1$  sans molécules de  $A$  en réactant). On définit alors la fonction  $g(N'_1) = R \cup \{A \rightarrow s\} \cup R'' \mid s''$ . Notons qu'en toute rigueur, la fonction  $g$  dépend des ensembles  $R$  et  $R'$ , mais on les suppose fixé et on omet donc de les mentionner.



Les conditions 1 et 5 du Lem. 4 sont vérifiées en appliquant récursivement la réaction  $A \rightarrow s$ . Les conditions 2, 6 et 7 sont également trivialement vérifiées. Pour la condition 3, on a  $N_1 \rightarrow N_2$  implique  $f(N_1) \xrightarrow{\epsilon} f(N_2)$ . En effet, si la réaction utilisée est  $A \rightarrow s$ , alors  $f(N_1) = f(N_2)$ . Sinon, on applique la réaction correspondante dans  $f(N_1)$ . Par récurrence, la condition est vérifiée. Réciproquement, si  $N'_1 \rightarrow N'_2$ , alors  $g(N'_1) \xrightarrow{+} g(N'_2)$ . En effet, il suffit d'appliquer la réaction correspondante dans  $g(N'_1)$  (qui n'a pas besoin de  $A$  d'après les conditions de la règle de simplification), puis de transformer toutes les molécules de  $A$  en  $s$  avec la réaction  $A \rightarrow s$ . Par récurrence, la condition 4 est vérifiée, et donc on peut appliquer le Lem. 4. Donc  $N \mid N'' \simeq_{\Omega} N' \mid N''$  et  $N \sim_{\Omega, I} N'$ .

**(INTERMÉDIAIRE - RÉACTANT)** : Ce cas est légèrement différent des autres. On ne peut pas appliquer le Lem. 4, car la condition 3 n'est valable que pour les réseaux ne possédant pas de molécules de l'espèce intermédiaire  $A$ . Il faut donc raisonner un peu autrement, de façon similaire à ce qui avait été fait pour la preuve d'exactitude de la règle **(CASCADE)**. On considère uniquement le cas où l'ensemble des réactions produisant (resp. consommant) l'espèce intermédiaire est non vide. Dans les autres cas, la règle est trivialement similaire soit à **(RÉACTIONS INAPPLICABLES)** suivie de **(ESPÈCE INUTILE)** (si aucune réaction ne produit l'espèce intermédiaire), ou directement de **(ESPÈCE INUTILE)** (si il n'y a pas de réaction de consommation).

Tout réseau  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  est de forme  $N_1 = R \cup R' \cup R'' \mid a_1 A + s''$ , avec  $R$  les réactions produisant  $A$  et  $R'$  celles consommant  $A$ . On définit la fonction  $f$  par  $f(N_1) = (R \triangleright_A R') \cup R'' \mid s''$ . Soit  $k$  un entier, et  $N_0, N_k \in \{(N \mid N'') \rightarrow^*\}$  deux réseaux accessibles, tels que  $A$  ne soit pas dans les solutions de ces deux réseaux. On veut montrer que  $N_0 \rightarrow^* N_k$  implique  $f(N_0) \rightarrow^* f(N_k)$ . Remarquons que pour tout  $N_1 \xrightarrow{r} N_2$  avec  $r \in R'$ , si  $A \in s(N_1)$  alors la réaction  $r$  peut commuter à gauche, c'est-à-dire qu'on peut également avoir la séquence  $N_1 \xrightarrow{r'} N_2$ . On peut alors décomposer et réorganiser la séquence  $N_0 \rightarrow^* N_k$  et obtenir une séquence de forme :

$$\begin{aligned} N_0 & \xrightarrow{r_0} N'_0 \xrightarrow{R''} N_1 \\ & \xrightarrow{r_1} N'_1 \xrightarrow{R''} N_2 \\ & \dots \\ & \xrightarrow{r_j} N'_j \xrightarrow{R''} N_{j+1}, \end{aligned}$$

avec  $N_{j+1} = N_k$ , et pour tout  $0 \leq i \leq j$ ,  $r_i \in R$ , et tous les réseaux dans les séquences  $N'_i \xrightarrow{R''} N_{i+1}$  ne contiennent pas de molécules de  $A$  dans leurs solutions. En d'autres termes, on applique une réaction de  $R$ , qui produit du  $A$ , puis on consomme toutes les molécules de  $A$  avec les réactions de  $R'$ , puis on applique une série de réactions du contexte n'utilisant pas  $A$ , et on recommence le même

schéma. On a alors trivialement  $f(N_i) \rightarrow_{r_i \triangleright_A R'} f(N'_i)$ , et  $f(N'_i) \rightarrow_{R(M_1)}^* f(N_{i+1})$ . Donc  $f(N_0) \rightarrow^* f(N_k)$ .

Réciproquement, tout réseau  $N'_1 \in \{(N' \mid N'') \rightarrow^*\}$  est de forme  $N'_1 = (R \triangleright_A R') \cup R'' \mid s''$ , on définit alors  $g(N'_1) = R \cup R' \cup R'' \mid s''$ . On a trivialement  $N'_1 \rightarrow^* N'_2$  implique  $g(N'_1) \rightarrow^* g(N'_2)$ .

Notons de plus que les fonctions  $f$  et  $g$  préservent l'observation, que pour tout  $N'_1 \in \{(N' \mid N'') \rightarrow^*\}$ , on a  $N'_1 = f(g(N'_1))$ , que pour tout  $N_1 \in \{(N \mid N'') \rightarrow^*\}$ , si  $A \notin s(N_1)$ , alors  $N_1 = g(f(N_1))$ , et que  $f(N) = N'$  et  $g(N') = N$ .

Soit  $\mathcal{A} \subseteq \{(N \mid N'') \rightarrow^*\}$  un  $\Omega$ -attracteur accessible, et  $N_1 \in \mathcal{A}$ . Comme  $R' \neq \emptyset$ , on peut appliquer une réaction de  $R'$ , et atteindre un réseau  $N_2 \in \{N_1 \rightarrow^*\}$  avec  $A \notin s(N_2)$ . Par la Prop. 2 (point 4),  $\mathcal{A}'' = \{N_2 \rightarrow^*\}$  est un  $\Omega$ -attracteur, avec la même observation que  $\mathcal{A}$ . On veut montrer que l'ensemble de réseaux  $\mathcal{A}' = \{f(N_2) \rightarrow^*\}$  est également un attracteur avec la même observation. Il est trivialement fermé, et accessible depuis  $N' \mid M$ . Soit  $N'_1 \in \mathcal{A}'$  un réseau, et  $o \in \Omega(\mathcal{A}')$  une observation. On veut montrer que  $o \in \Omega(\{N'_1 \rightarrow^*\})$ . Il y a donc un réseau  $N'_2 \in \mathcal{A}'$  avec  $o = \Omega(N'_2)$ . On a  $f(N_2) \rightarrow^* N'_2$ , donc  $N_2 = g(f(N_2)) \rightarrow^* g(N'_2)$ , et comme  $\Omega(g(N'_2)) = \Omega(N'_2)$ , on a  $o \in \Omega(\mathcal{A}'')$ . De façon similaire,  $g(N'_1) \in \mathcal{A}''$ . Donc par  $\Omega$ -connectivité, il existe  $N_3 \in \mathcal{A}''$  tel que  $g(N'_1) \rightarrow^* N_3$  et  $o = \Omega(N_3)$ . En appliquant une réaction de  $R'$ , il existe  $N_4 \in \{N_3 \rightarrow^*\} \subseteq \{g(N'_1) \rightarrow^*\}$  avec  $\Omega(N_4) = o$  et  $A \notin s(N_4)$ , donc  $N'_1 = f(g(N'_1)) \rightarrow^* f(N_4)$ , et donc  $o \in \Omega(\{N'_1 \rightarrow^*\})$ , donc  $\mathcal{A}'$  est bien un attracteur.

Réciproquement, si  $\mathcal{A}'$  est un attracteur accessible, et  $N'_1 \in \mathcal{A}'$ , alors  $\mathcal{A} = \{g(N'_1) \rightarrow^*\}$  est aussi un attracteur. En effet, il est trivialement fermé et accessible. Soit  $N_1 \in \mathcal{A}$  et  $o \in \Omega(\mathcal{A})$ . Il existe  $N_2 \in \{g(N'_1) \rightarrow^*\}$  avec  $o = \Omega(N_2)$ . En appliquant les réactions de  $R'$ , il existe  $N_3 \in \{N_2 \rightarrow^*\}$  avec  $o = \Omega(N_3)$  et  $A \notin s(N_3)$ . Donc  $f(N_3) \in \{f(g(N'_1)) \rightarrow^*\} = \{N'_1 \rightarrow^*\} \subseteq \mathcal{A}'$ . Donc  $o \in \Omega(\mathcal{A}')$ . De même, il existe  $N_4 \in \{N_1 \rightarrow^*\}$  avec  $A \notin s(N_4)$ . Donc  $f(N_4) \in \mathcal{A}'$ , et par  $\Omega$ -connectivité il existe  $N'_2 \in \{f(N_4) \rightarrow^*\}$  et  $o = \Omega(N'_2)$ . Donc  $g(N'_2) \in \{g(f(N_4)) \rightarrow^*\} = \{N_4 \rightarrow^*\} \subseteq \{N_1 \rightarrow^*\}$ . Donc  $o \in \Omega(\{N_1 \rightarrow^*\})$ , donc  $\mathcal{A}$  est bien un attracteur.

Les preuves pour les divergences sont similaires. Par conséquent  $N \sim_{\Omega} N'$ .

**(INTERMÉDIAIRE - PRODUIT)** : Ce cas est quasiment similaire au précédent.

Tout réseau  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  est de forme  $N_1 = R \cup R' \cup R'' \mid a_1 A + s''$  avec  $A \notin s''$ ,  $R$  les réactions produisant  $A$  et  $R'$  celles consommant  $A$ . On définit la fonction  $f$  par  $f(N_1) = (R \triangleleft_A R') \cup R'' \mid s''$ . Alors pour tout réseau  $N_0, N_k \in \{(N \mid N'') \rightarrow^*\}$ , si il n'y a pas de molécules de  $A$  dans les solutions de ces deux réseaux, alors  $N_0 \rightarrow^* N_k$  implique  $f(N_0) \rightarrow^* f(N_k)$ . La preuve est similaire au cas précédent, en remarquant que cette fois se sont les réactions de  $R$  qui peuvent commuter vers la droite, c'est-à-dire que pour  $N_1 \xrightarrow{r} N_2$  avec  $r \in R$  et  $A \in s(N_2)$ , la séquence  $N_1 \xrightarrow{r'} N_2$  est aussi valide.

Réciproquement, tout réseau  $N'_1 \in \{(N' \mid N'') \rightarrow^*\}$  est de forme  $N'_1 = (R \triangleleft_A R') \cup$

$R'' \mid s''$  avec  $A \notin s''$ , on définit alors  $g(N'_1) = R \cup R' \cup R'' \mid s''$ . On a trivialement  $N'_1 \rightarrow^* N'_2$  implique  $g(N'_1) \rightarrow^* g(N'_2)$ .

Notons de plus que les fonctions  $f$  et  $g$  préservent l'observation, que pour tout  $N'_1 \in \{(N' \mid N'') \rightarrow^*\}$ , on a  $N'_1 = f(g(N'_1))$ , que pour tout  $N_1 \in \{(N \mid N'') \rightarrow^*\}$ , si  $A \notin s(N_1)$ , alors  $N_1 = g(f(N_1))$ , et que  $f(N) = N'$  et  $g(N') = N$ .

Enfin, on veut montrer la propriété suivante : soit  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  sans molécule de  $A$  dans sa solution, et  $N_2 \in \{N_1 \rightarrow^*\}$ . Alors il existe  $N_3 \in \{N_1 \rightarrow^*\}$  sans molécule de  $A$  et tel que  $\Omega(N_2) = \Omega(N_3)$ . Notons que cette propriété était directement vérifiée dans la règle précédente (**INTERMÉDIAIRE - RÉACTANT**). Soit  $N_4$  un réseau tel que  $N_1 \rightarrow^* N_4 \rightarrow^k N_2$ , avec  $k \geq 0$ , sans molécule de  $A$ , mais que tout réseau dans la séquence  $N_4 \rightarrow^* N_2$  ait des molécules de  $A$  dans sa solution (sauf  $N_2$  et  $N_4$ ). Il existe bien un tel réseau puisque  $N_1$  n'a pas de molécule de  $A$ . On va montrer la propriété par récurrence sur la taille  $k$  de  $N_4$  à  $N_2$ . Si  $k = 0$ , alors  $N_2 = N_4$ , donc  $N_2$  n'a pas de molécule de  $A$ , on prend donc  $N_3 = N_2$ . Sinon, la réaction appliquée dans  $N_4$  est forcément une réaction de  $R$ , créant du  $A$ . On a en fait  $N_4 \xrightarrow{R} N_5$ , avec soit  $N_5 = N_2$ , soit  $N_5 \xrightarrow{r'} N_2$ , avec  $r' \notin R$ . Dans le premier cas,  $\Omega(N_4) = \Omega(N_2)$ , et donc la propriété est vraie. Dans le second cas, si  $r' \in R''$ , alors on peut la commuter complètement à gauche avec les réactions de  $R$ , on obtient donc une séquence  $N_4 \xrightarrow{r'} N'_4 \xrightarrow{R} N_5$ . La séquence  $N'_4 \rightarrow^{k-1} N_2$  est plus courte, et  $N'_4$  n'a pas de molécule de  $A$ , donc par récurrence la propriété est vraie. Enfin, si  $r' \in R'$ , on peut également la commuter à gauche jusqu'à atteindre une séquence de type  $N_4 \xrightarrow{R} N'_4 \xrightarrow{r'} N_5$ , avec  $N'_4$  sans molécule de  $A$ . Encore une fois, la séquence partant de  $N'_4$  est plus petite, donc la propriété est vraie par récurrence.

Soit  $\mathcal{A} \subseteq \{(N \mid N'') \rightarrow^*\}$  un  $\Omega$ -attracteur accessible, et  $N_1 \in \mathcal{A}$ . Comme  $R' \neq \emptyset$  d'après les conditions de la règle, donc on peut appliquer une réaction de  $R'$ , et atteindre un réseau  $N_2 \in \{N_1 \rightarrow^*\}$  avec  $A \notin s(N_2)$ . Par la Prop. 2 (point 4),  $\mathcal{A}'' = \{N_2 \rightarrow^*\}$  est un  $\Omega$ -attracteur, avec la même observation que  $\mathcal{A}$ . On veut montrer que l'ensemble de réseaux  $\mathcal{A}' = \{f(N_2) \rightarrow^*\}$  est également un attracteur avec la même observation. Il est trivialement fermé, et accessible depuis  $N' \mid M$ . Soit  $N'_1 \in \mathcal{A}'$  un réseau, et  $o \in \Omega(\mathcal{A}')$  une observation. On veut montrer que  $o \in \Omega(\{N'_1 \rightarrow^*\})$ . Il y a donc un réseau  $N'_2 \in \mathcal{A}'$  avec  $o = \Omega(N'_2)$ . On a  $f(N_2) \rightarrow^* N'_2$ , donc  $N_2 = g(f(N_2)) \rightarrow^* g(N'_2)$ , et comme  $\Omega(g(N'_2)) = \Omega(N'_2)$ , on a  $o \in \Omega(\mathcal{A}'')$ . De façon similaire,  $g(N'_1) \in \mathcal{A}''$ . Donc par  $\Omega$ -connectivité, il existe  $N_3 \in \mathcal{A}''$  tel que  $g(N'_1) \rightarrow^* N_3$  et  $o = \Omega(N_3)$ . Par la propriété supplémentaire que l'on a prouvé ci-dessus, il existe un réseau  $N_4 \in \{g(N'_1) \rightarrow^*\}$ , avec  $o = \Omega(N_4)$  et  $A \notin s(N_4)$ . Donc  $N'_1 = f(g(N'_1)) \rightarrow^* f(N_4)$ , d'où  $o \in \Omega(\{N'_1 \rightarrow^*\})$ , et  $\mathcal{A}'$  est bien un attracteur.

Réciproquement, si  $\mathcal{A}'$  est un attracteur accessible, et  $N'_1 \in \mathcal{A}'$ , alors  $\mathcal{A} = \{g(N'_1) \rightarrow^*\}$  est aussi un attracteur. En effet, il est trivialement fermé et accessible. Soit  $N_1 \in \mathcal{A}$  et  $o \in \Omega(\mathcal{A})$ . Il existe  $N_2 \in \{g(N'_1) \rightarrow^*\}$  avec  $o = \Omega(N_2)$ . Avec la propriété ci-dessus, il existe  $N_3 \in \{g(N'_1) \rightarrow^*\}$  avec  $o = \Omega(N_3)$  et  $A \notin s(N_3)$ . Donc

$f(N_3) \in \{f(g(N'_1)) \rightarrow^*\} = \{N'_1 \rightarrow^*\} \subseteq \mathcal{A}'$ . Donc  $o \in \Omega(\mathcal{A}')$ . De même, il existe  $N_4 \in \{N_1 \rightarrow^*\}$  avec  $A \notin s(N_4)$ . Donc  $f(N_4) \in \mathcal{A}'$ , et par  $\Omega$ -connectivité il existe  $N'_2 \in \{f(N_4) \rightarrow^*\}$  et  $o = \Omega(N'_2)$ . Donc  $g(N'_2) \in \{g(f(N_4)) \rightarrow^*\} = \{N_4 \rightarrow^*\} \subseteq \{N_1 \rightarrow^*\}$ . Donc  $o \in \Omega(\{N_1 \rightarrow^*\})$ , donc  $\mathcal{A}$  est bien un attracteur.

Les preuves pour les divergences sont similaires. Par conséquent  $N \sim_{\Omega, \mathcal{I}} N'$ .

**(RÉDUCTION INVERSIBLE)** : Cette preuve est basée sur le Lem. 4.

On pose  $N = R \mid s'$  le réseau initial. Tout réseau  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  étant de forme  $N_1 = R \cup R'' \mid a_1A + s''$ , avec  $A \notin s''$ , on définit  $f(N_1) = R[A := s] \cup R'' \mid a_1s + s''$ . Réciproquement, tout réseau  $N_1 \in \{(N' \mid N'') \rightarrow^*\}$  est de forme  $N_1 = R[A := s] \cup R'' \mid s''$ , on définit alors  $g(N'_1) = R \cup R'' \mid s''$ .

On a directement  $N_1 \rightarrow_r N_2$  implique  $f(N_1) \rightarrow_{r[A:=s]} f(N_2)$ , donc par récurrence on vérifie la condition 3 du Lem. 4. De même, si  $N'_1 \rightarrow_{r'} N'_2$ , alors il existe  $r \in R \cup R'$  telle que  $r[A := s] = r'$ . On peut donc transformer du  $s$  en  $A$ , utiliser la réaction  $r$ , puis retransformer les molécules de  $A$  en  $s$ . Par récurrence, on vérifie la condition 4. En transformant toutes les molécules de  $A$  en  $s$ , on vérifie la condition 5. Comme  $f(g(N'_1)) = N'_1$ , la condition 6 est directement vérifiée. Puisque  $\Omega(A) = \Omega(s)$ , on vérifie aussi la condition 7. Enfin, les conditions 1 et 2 sont trivialement vraies aussi. On peut donc appliquer le Lem. 4, et  $N \sim_{\Omega, \mathcal{I}} N'$ .

**(DÉCOMPOSITION - SUPPRESSION)** : Cette preuve est basée sur le Lem. 4.

Tout  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  est de forme  $N_1 = \{s_1 \rightarrow s_2\} \cup R \cup R'' \mid s''$ , avec  $R$  un ensemble de réactions pouvant imiter la réaction  $s_1 \rightarrow s_2$ . On définit  $f(N_1) = R \cup R'' \mid s''$ , et réciproquement  $g = f^{-1}$ . On a trivialement  $N_1 \rightarrow N_2$  implique  $f(N_1) \rightarrow^* f(N_2)$  d'après les conditions de la règle, et  $N'_1 \rightarrow N'_2$  implique  $f(N'_1) \rightarrow f(N'_2)$ . Donc les conditions 3 et 4 du Lem. 4 sont vérifiées. Les autres conditions le sont également trivialement, et donc  $N \sim_{\Omega, \mathcal{I}} N'$ .

**(DÉCOMPOSITION - SIMPLIFICATION)** : Pour tout  $N_1 \in \{(N \mid N'') \rightarrow^*\}$ , de forme  $N_1 = \{s_1 \rightarrow s_2\} \cup R \cup R'' \mid s''$ , on définit  $f(N_1) = \{s'_1 \rightarrow s'_2\} \cup R \cup R'' \mid s''$ , et réciproquement  $g = f^{-1}$ . Toutes les conditions du Lem. 4 sont alors trivialement vérifiées, et  $N \sim_{\Omega, \mathcal{I}} N'$ .

**(SYMÉTRIE - RÉACTANTS)** On va cette fois utiliser le Lem. 5.

Tout réseau  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  est de forme  $N_1 = R \cup R' \cup R'' \mid a_1A_1 + a_2A_2 + s''$  avec  $A_1, A_2 \notin s''$ , avec  $R$  des réactions où  $A_1$  et  $A_2$  sont réactants-symétriques, et  $R'$  des réactions transformant  $A_1$  en  $A_2$  ou réciproquement. Pour tout réseau de cette forme, on définit la fonction  $f(N_1) = R[A_2 := A_1] \cup R'' \mid (a_1 + a_2)A_1 + s''$ . On a  $f(N) = N'$  (condition 2). On va montrer que pour toute paire de réseaux

$N_1, N_2 \in \{(N \mid N'') \rightarrow^*\}$ , on a :

$$N_1 \rightarrow N_2 \quad \text{implique} \quad f(N_1) \rightarrow^\epsilon f(N_2).$$

Soit  $r$  la réaction telle que  $N_1 \rightarrow_r N_2$ . Si  $A_2 \notin r$ , alors directement  $f(N_1) \rightarrow_r f(N_2)$ . Si  $A_2 \in r$ , alors soit  $r \in R$  soit  $r \in R'$ . Si  $r \in R$ , alors  $r' = r[A_2 := A_1] \in R[A_2 := A_1]$  est applicable sur  $f(N_1)$ , et on a bien  $f(N_1) \rightarrow_{r'} f(N_2)$ . Si  $r \in R'$ , alors  $f(N_1) = f(N_2)$ . Par récurrence, on a la condition 3 du lemme :

$$N_1 \rightarrow^* N_2 \quad \text{implique} \quad f(N_1) \rightarrow^* f(N_2).$$

Réciproquement, tout réseau  $N'_1 \in \{(N' \mid N'') \rightarrow^*\}$  est de forme  $N'_1 = R[A_2 := A_1] \cup R'' \mid a'A_1 + s''$ , avec  $A_1, A_2 \notin s''$ . On définit l'ensemble des réseaux  $g(N'_1) = \{(R \cup R' \cup R'' \mid a_1A_1 + a_2A_2 + s'' \mid a' = a_1 + a_2)\}$ . On a  $N \in g(N')$  (condition 1). On va montrer que pour toute paire de réseaux  $N'_1, N'_2 \in \{(N' \mid N'') \rightarrow^*\}$ , on a :

$$N'_1 \rightarrow N'_2 \quad \text{implique} \quad \forall N_1 \in g(N'_1). \exists N_2 \in g(N'_2). N_1 \rightarrow^* N_2.$$

Soit  $r' = s'_1 \rightarrow s'_2$  la réaction telle que  $N'_1 \rightarrow_{r'} N'_2$ . Si  $r' \in R(M_1)$ , alors  $A_2 \notin r'$ , donc  $r'$  est applicable dans  $N_1$ . Le réseau  $N_2$  obtenu appartient bien à  $g(N'_2)$ , donc  $N_1 \rightarrow_{r'} N_2$ . Sinon,  $r' \in R[A_2 := A_1]$ , donc il existe  $r = s_1 \rightarrow s_2 \in R$  avec  $r[A_2 := A_1] = r'$ . Donc en particulier  $A_1 \xrightarrow[s_1, s'_1]{sol} A_2$ . On a également  $A_1 \xrightarrow[s(N_1), s(N'_1)]{sol} A_2$ .

Il existe alors  $s''_1 \subseteq s(N_1)$  avec  $A_1 \xrightarrow[s''_1, s'_1]{sol} A_2$ , et donc  $A_1 \xrightarrow[s'_1, s_1]{sol} A_2$ . Comme  $A_1 \xrightarrow[R]{rea} A_2$ , on a alors  $(s''_1 \mid R) \rightarrow^*(s''_2 \mid R)$ , avec  $s''_2$  telle que  $A_1 \xrightarrow[s''_2, s_2]{sol} A_2$ . Mais alors  $A_1 \xrightarrow[s'_1, s_1]{sol} A_2$ , et donc  $N_1 \rightarrow^* N_2$  avec  $N_2 \in g(N'_2)$ . Par récurrence immédiate, on a la condition 4 du lemme :

$$N'_1 \rightarrow^* N'_2 \quad \text{implique} \quad \forall N_1 \in g(N'_1). \exists N_2 \in g(N'_2). N_1 \rightarrow^* N_2.$$

De plus, trivialement,  $\Omega(f(N_1)) = \Omega(N_1)$  et pour tout  $N_1 \in g(N'_1)$ ,  $\Omega(N_1) = \Omega(N'_1)$  (condition 7). On a aussi  $N_1 \in g(f(N_1))$  et  $f(g(N'_1)) = \{N'_1\}$  (conditions 5 et 6). Donc les fonctions  $f$  et  $g$  vérifient les propriétés du Lem. 5, donc les deux réseaux sont attracteurs-congruents.

**(SYMÉTRIE - PRODUITS)** La preuve est légèrement similaire à celle du Lem. 5, mais la condition 4 n'est vérifiée que pour les réseaux n'ayant pas de molécules de  $A_1$  dans leurs solutions initiales.

Tout réseau accessible depuis le réseau initial dans son contexte est de forme  $N_1 = R \cup R' \cup R'' \mid a_1A_1 + a_2A_2 + s''$ , avec  $A_1, A_2 \notin s''$ ,  $R$  des réactions où  $A_1$  et  $A_2$  sont produits-symétriques et  $R'$  des réactions transformant  $A_1$  en  $A_2$  ou

réciroquement. On définit  $f(N_1) = R[[A_2 := A_1] \cup R'' \mid (a_1 + a_2)A_1 + s'']$ . Alors comme ci-dessus on a directement,  $N_1 \rightarrow^* N_2$  implique  $f(N_1) \rightarrow^* f(N_2)$ . De plus,  $f(N) = N'$ , et  $\Omega(f(N_1)) = \Omega(N_1)$ .

Réciroquement, pour tout réseau  $N'_1 = R[A_2 := A_1] \cup R'' \mid a'A_1 + s''$  accessible depuis  $N' \mid N''$ , on définit  $g(N'_1) = \{N_1 \mid f(N_1) = N'_1\}$ . On a  $N \in g(N')$  et pour tout  $N_1 \in g(N'_1)$ ,  $\Omega(N_1) = \Omega(N'_1)$ . On va montrer que pour tout  $N'_2, N'_1 \in \{(N' \mid N'') \rightarrow^*\}$ , on a :

$$N'_1 \rightarrow N'_2 \quad \text{implique} \quad \forall N_2 \in g(N'_2). \exists N_1 \in g(N'_1). N_1 \rightarrow^* N_2. \quad (3.9)$$

Notons que c'est la condition inverse de celle du Lem. 5. Supposons que  $N'_1 \rightarrow_{r'} N'_2$  avec  $r' = s'_1 \rightarrow s'_2$ , et soit  $N_2 \in g(N'_2)$ . Si  $r' \in R''$ , trivialement  $N_1 \rightarrow_{r'} N_2$  avec  $N_1 \in g(N'_1)$ . Si  $r' \in R[A_2 := A_1]$ , il existe  $r = s_1 \rightarrow s_2 \in R$  avec  $r[A_2 := A_1] = r'$ .

En particulier on a  $A_1 \xrightarrow[s_1, s'_1]{sol} A_2$  et  $A_1 \xrightarrow[s_2, s'_2]{sol} A_2$ .

$$N_2 \in g(N'_2)$$

implique  $A_1 \xrightarrow[s(N_2), s(N'_2)]{sol} A_2$ , donc il existe  $s''_1 \subseteq s(N_2)$  avec  $A_1 \xrightarrow[s''_1, s'_2]{sol} A_2$  et donc

$A_1 \xrightarrow[s''_1, s'_2]{sol} A_2$ . Comme  $A_1 \xrightarrow[R]{pro} A_2$ , on a alors  $(s''_1 \mid R) \rightarrow^* (s'_2 \mid R)$ , avec  $s''_1$  telle

que  $A_1 \xrightarrow[s''_1, s'_2]{sol} A_2$ . Mais alors  $A_1 \xrightarrow[s''_1, s'_1]{sol} A_2$ , et donc  $N_1 \rightarrow^* N_2$  avec  $N_1 \in g(N'_1)$ . Par récurrence immédiate, on a :

$$N'_1 \rightarrow^* N'_2 \quad \text{implique} \quad \forall N_2 \in g(N'_2). \exists N_1 \in g(N'_1). N_1 \rightarrow^* N_2.$$

On considère deux réseaux  $N'_1, N'_2$  sans molécule de  $A_2$  dans  $N'_1$ . On a alors :

$$N'_1 \rightarrow^* N'_2 \quad \text{implique} \quad \forall N_1 \in g(N'_1). \forall N_2 \in g(N'_2). N_1 \rightarrow^* N_2. \quad (3.10)$$

En effet, d'après la propriété précédente, il existe un  $N_1 \in g(N'_1)$  qui vérifie cette condition. Mais  $N'_1$  n'a pas de molécule de  $A_2$ , ce qui implique que  $|g(N'_1)| = 1$ , c'est-à-dire  $g(N'_1) = \{N_1\}$ , et donc la condition est vérifiée.

Soit  $\mathcal{A}$  un  $\Omega$ -attracteur accessible depuis  $N \mid N''$ . Soit  $N_1 \in \mathcal{A}$ . D'après les conditions de la règle, il existe des réactions  $r_1 = A_1 \rightarrow s_1$  et  $r_2 = A_2 \rightarrow s_2$  dans  $R(N_1)$ . Donc en les appliquant, il existe  $N_2 \in \mathcal{A}$  avec  $A_1, A_2 \notin s(N_2)$ . Soit  $N'_2 = f(N_2)$  et  $\mathcal{A}' = \{N'_2 \rightarrow^*\}$ . Remarquons que  $A_1 \notin s(N'_2)$ .  $(N \mid N'') \rightarrow^* N_2$  implique  $(N' \mid N'') \rightarrow^* N'_2$  (3.9), donc  $\mathcal{A}'$  est accessible. Il est de plus fermé. Soit  $N'_3 \in \mathcal{A}'$ . On a  $N'_2 \rightarrow^* N'_3$  implique qu'il existe  $N_3 \in g(N'_3)$  (en fait pour tout  $N_3 \in g(N'_3)$ ) tel que  $N_2 \rightarrow^* N_3$  (3.10), donc  $N_3 \in \mathcal{A}$ . Soit  $o \in \Omega(\mathcal{A}')$ , c'est-à-dire qu'il existe  $N'_4 \in \mathcal{A}'$  avec  $o = \Omega(N'_4)$ . Comme précédemment, il existe  $N_4 \in g(N'_4) \cap \mathcal{A}$ . Donc  $o = \Omega(N_4) \in \Omega_{(attr)}$ . Par  $\Omega$ -connectivité, il existe  $N_5$  tel

que  $N_3 \rightarrow^* N_5$  et  $o = \Omega(N_5)$ , mais alors  $N'_3 = f(N_3) \rightarrow^* f(N_5)$  (3.9) et  $O \in \{N'_3 \rightarrow^*\}$ . Donc  $\mathcal{A}'$  est  $\Omega$ -connecté, donc c'est un attracteur.

Réciproquement, soit  $\mathcal{A}'$  un  $\Omega$ -attracteur accessible depuis  $N' \mid N''$ . Soit  $N'_1 \in \mathcal{A}'$ . Il existe  $r' = A_1 \rightarrow s \in R(N_1)$  d'après les conditions de la règle. Donc il existe  $N'_2 \in \mathcal{A}'$  avec  $A_1 \notin s(N'_2)$ . Soit  $N_2$  l'unique réseau tel que  $N_2 \in g(N'_2)$ , et  $\mathcal{A} = \{N_2 \rightarrow^*\}$ .  $\mathcal{A}$  est fermé, et accessible, puisque  $N' \mid N'' \rightarrow^* N'_2$  implique  $N \mid N'' \rightarrow^* N_2$  par 3.9. Soit  $N_3 \in \mathcal{A}$ , et  $N_4 \in \{N_3 \rightarrow^*\}$  avec  $A_1, A_2 \notin s(N_4)$ . Alors  $N'_4 = f(N_4) \in \mathcal{A}'$  par 3.9. Soit  $o \in \Omega(\mathcal{A})$ , alors  $o \in \Omega(\mathcal{A}')$ . Par  $\Omega$ -connectivité, il existe  $N'_5 \in \{N'_4 \rightarrow^*\}$  avec  $o = \Omega(N'_5)$ . D'après 3.10, on a alors  $N_4 \rightarrow^* N_5$ , avec  $N_5 \in g(N'_4)$ . Donc  $N_3 \rightarrow^* N_5$ , et  $o \in \Omega(\{N_3 \rightarrow^*\})$ . Donc  $\mathcal{A}$  est un attracteur.

Trivialement  $\mathcal{A}$  et  $\mathcal{A}'$  ont la même observation, donc  $N \mid N''$  et  $N' \mid N''$  vérifient les mêmes prédicats de convergence (la preuve pour la divergence est similaire). Donc  $N \sim_{\Omega, \mathcal{I}} N'$ .

**(SYMÉTRIE - IMITATION)** : On utilise une variante du Lem. 4, où la condition 5 est remplacée par une condition de type  $g(f(N_1)) = h(N_1)$ .

Pour tout  $N_1 = R \cup R'' \mid aA + s''$ , on définit  $f(N_1) = N_1[A := s]$ . On a directement  $N_1 \rightarrow^* N_2$  implique  $f(N_1) \rightarrow^* f(N_2)$ . De plus  $\Omega(f(N_1)) = \Omega(N_1)$  et  $f(N) = N'$ .

On définit également la fonction  $h$  sur tout réseau  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  par  $h(N_1) = R \cup R'' \mid as + s''$ . Alors la condition  $A \underset{R}{\overset{app}{\rightarrow}} s$  implique directement que :

$$N_1 \rightarrow^* N_2 \quad \text{implique} \quad h(N_1) \rightarrow^* h(N_2).$$

En d'autres termes, pour toute séquence de réductions, on a une séquence équivalente qui utilise  $s$  à la place de  $A$ . On a aussi  $\Omega(h(N_1)) = \Omega(N_1)$  et  $h(N) = N$ .

Réciproquement, pour tout  $N'_1 = R[A := s] \cup R'' \mid s''$ , on définit  $g(N'_1) = h(f^{-1}(N'_1))$ . Donc  $g(N'_1)$  est l'unique réseau  $N_1$  tel que  $f(N_1) = N'_1$  et  $A \notin s(N_1)$ . Alors on a trivialement :

$$N'_1 \rightarrow^* N'_2 \quad \text{implique} \quad g(N_1) \rightarrow^* g(N_2).$$

De plus,  $\Omega(g(N'_1)) = \Omega(N'_1)$  et  $g(N') = N$ . Enfin,  $g(f(N_1)) = h(N_1)$  et  $f(g(N'_1)) = N'_1$ .

Soit  $\mathcal{A}$  un attracteur accessible depuis  $(N \mid N'')$  et  $N_1 \in \mathcal{A}$ . On applique récursivement la réaction  $A \rightarrow s' \in R$ , jusqu'à obtenir un réseau  $N_2 \in \mathcal{A}$  avec  $A \notin s(N_2)$ . On pose  $\mathcal{A}'' = \{N_2 \rightarrow^*\}$ . C'est un attracteur de même observation que  $\mathcal{A}$  d'après les propriétés de la Prop. 2. Soit  $N'_2 = f(N_2)$ . Remarquons qu'on a  $N_2 = g(N'_2)$ . On pose  $\mathcal{A}' = \{N'_2 \rightarrow^*\}$ . C'est un ensemble fermé et accessible depuis  $N' \mid N''$ . Soit  $N'_3 \in \mathcal{A}'$ , donc  $N'_2 \rightarrow^* N'_3$ . Alors  $N_2 = g(N'_2) \rightarrow^* g(N'_3)$ , donc  $g(N'_3) \in \mathcal{A}''$ . De même, si  $o \in \Omega(\mathcal{A}')$ , alors  $o \in \Omega(\mathcal{A}'')$ . Par  $\Omega$ -connectivité, il existe

$N_4 \in \{g(N_3) \rightarrow^*\}$  avec  $o = \Omega(N_3)$ , mais alors  $f(N_4) \in \{f(g(N_3)) \rightarrow^*\} = \{N_3 \rightarrow^*\}$ . Donc  $\mathcal{A}'$  est  $\Omega$ -connecté, et  $c'$  est un attracteur.

Réciproquement, soit  $\mathcal{A}'$  un attracteur accessible depuis  $(N \mid N'')$ , et  $N'_1 \in \mathcal{A}'$ . Alors  $\mathcal{A} = \{g(N'_1) \rightarrow^*\}$  est fermé et accessible. Soit  $N_2 \in \mathcal{A}$ , on applique récursivement  $A \rightarrow s' \in R$  jusqu'à atteindre  $N_3 \in \mathcal{A}$  avec  $A \notin s(N_2)$ . On a donc  $g(f(N_3)) = h(N_3)$ . On a  $N_3 \in \mathcal{A}$ , donc  $f(N_3) \in \mathcal{A}'$ . De même, pour tout  $o \in \Omega(\mathcal{A})$ , on a aussi  $o \in \Omega(\mathcal{A}')$ . Donc par connectivité, il existe  $N'_4$  tel que  $f(N_3) \rightarrow^* N'_4$  et  $o = \Omega(N'_4)$ . Mais on a donc  $N_3 = g(f(N_3)) \rightarrow^* g(N'_4)$ . Donc  $o \in \{N_3 \rightarrow^*\} \subseteq \{N_2 \rightarrow^*\}$ . Donc  $\mathcal{A}$  est un attracteur.

Les attracteurs  $\mathcal{A}$  et  $\mathcal{A}'$  ont trivialement la même observation. Les preuves pour la divergence sont similaires. Donc  $N$  et  $N'$  sont attracteurs-congruents.

**(RÉACTIONS INAPPLICABLES)** : Cette preuve utilise sur le Lem. 4.

Tout  $N_1 \in \{(N \mid N'') \rightarrow^*\}$  est de forme  $N_1 = R_1 \cup R_2 \cup R'' \mid a_1 A + s''$ , avec  $R_1$  les réactions inapplicables nécessitant des molécules de  $A$ , et  $R_2$  des réactions applicables mais ne pouvant pas augmenter la quantité de  $A$ . On définit  $f(N_1) = R_2 \cup R'' \mid a_1 A + s''$ , et  $g = f^{-1}$ . Par récurrence immédiate, pour tout  $N_1$ ,  $a_1 \leq a_0$ . Donc pour tout  $r = s_1 \rightarrow s_2 \in R_1$ , on a  $s_1(A) > a_0 \geq a_1$ , donc  $r$  n'est pas applicable dans  $N_1$ . Donc trivialement,  $N_1 \rightarrow N_2$  si et seulement si  $f(N_1) \rightarrow f(N_2)$ . Toutes les conditions du Lem. 4 sont donc vérifiées, et  $N \sim_{\Omega, I} N'$ .

**(PRODUCTIONS GROUPÉES)** : Pour cette règle, on va appliquer une variante du Lem. 5.

Soit  $V = \text{Vars}(R)$  l'ensemble des espèces pouvant être produites. Pour tout  $N_1 \in \{(N \mid N'') \rightarrow^*\}$ , de forme  $R \cup R'' \mid s''$ , on définit l'ensemble (infini) des réseaux  $f(N_1) = \{(\emptyset \rightarrow s) \cup R'' \mid s'' + s'_1 \mid \text{Vars}(s'_1) = V\}$ . On peut donc ajouter à la solution de  $N_1$  un nombre arbitraire de molécules des espèces de  $V$ . De même, pour  $N'_1 = \{\emptyset \rightarrow s\} \cup R'' \mid s'' \in \{(N' \mid N'') \rightarrow^*\}$ , on définit  $g(N'_1) = \{R \cup R'' \mid s'' + s_1 \mid \text{Vars}(s_1) = V\}$ . Alors, pour tout  $N_1, N_2 \in \{(N \mid N'') \rightarrow^*\}$ , on prouve que :

$$N_1 \rightarrow^* N_2 \quad \text{implique} \quad \forall N'_1 \in f(N_1). \exists N'_2 \in f(N_2). N'_1 \rightarrow^* N'_2.$$

Il suffit en effet de remplacer chaque occurrence d'une réaction  $\emptyset \rightarrow s_i \in R$  par  $|s_i|$  occurrences de la réaction  $\emptyset \rightarrow s$ . Réciproquement, pour tout  $N'_1, N'_2 \in \{(N' \mid N'') \rightarrow^*\}$ , on a :

$$N'_1 \rightarrow^* N'_2 \quad \text{implique} \quad \forall N_1 \in g(N'_1). \exists N_2 \in g(N'_2). N_1 \rightarrow^* N_2.$$

On remplace en effet toute occurrence de  $\emptyset \rightarrow s$  par une occurrence de chaque réaction  $r_i \in R$ .

Pour tout  $N_1 \in \{(N \mid N'') \rightarrow^*\}$ , on a également :

$$\forall N_2 \in g(f(N_1)). \forall N_3 \in \{N_2 \rightarrow^*\}. \exists N_4 \in g(f(N_3)). N_1 \rightarrow^* N_4.$$



Le cas pour les réseaux de  $\{(N' \mid N'') \rightarrow^*\}$  est similaire.

De plus, ces fonctions préservent l'observation : pour tout réseau  $N'_1 \in f(N_1)$ , on a  $\Omega(N'_1) = \Omega(N_1)$ , et réciproquement  $\forall N_1 \in g(N'_1)$ , on a  $\Omega(N'_1) = \Omega(N_1)$ .

Pour tout  $\Omega$ -attracteur  $\mathcal{A} \subseteq \{(N \mid N'') \rightarrow^*\}$ , soit  $N_1 \in \mathcal{A}$ , et  $N'_1 \in f(N_1)$  tel que  $N'_1 \in \{(N' \mid N'') \rightarrow^*\}$ . Soit  $\mathcal{A}' = \{N'_1 \rightarrow^*\}$  un ensemble fermé et accessible de réseaux. Soit  $N'_2 \in \{N'_1 \rightarrow^*\}$ , alors il existe  $N_2 \in g(N'_2) \cap \{N_1 \rightarrow^*\}$ , donc  $N_2 \in \mathcal{A}$ . Soit  $o \in \Omega(\mathcal{A}')$ , il existe  $N'_3 \in \{N'_1 \rightarrow^*\}$  avec  $\Omega(N'_3) = o$  et  $N_3 \in g(N'_3) \cap \{N_1 \rightarrow^*\}$  avec  $\Omega(N_3) = \Omega(N'_3)$ . Donc  $o \in \Omega(\mathcal{A}) \subseteq \Omega(N_2)$  par  $\Omega$ -connectivité. Donc il existe  $N_4 \in \{N_2 \rightarrow^*\}$  avec  $\Omega(N_4) = o$ . Donc il existe  $N'_4 = g(N_4) \cap \{g(N_2) \rightarrow^*\}$ , c'est-à-dire  $N'_4 \in \{g(f(N'_2)) \rightarrow^*\}$ , avec toujours  $\Omega(N'_4) = o$ . Donc il existe aussi  $N'_5 \in \{N'_2 \rightarrow^*\}$  avec  $\Omega(N'_5) = o$ , et donc  $\mathcal{A}'$  est  $\Omega$ -connecté, donc c'est un attracteur.

La réciproque, ainsi que les preuves de divergence, sont similaires, et donc  $N \sim_{\Omega, \mathcal{I}} N'$ .

**(SOLUTION INDIVISIBLE)** : Pour tout  $N_1 = R \cup R'' \mid s_1 \in \{(N \mid N'') \rightarrow^*\}$ , on définit  $f(N_1) = R[s := A] \cup R'' \mid s_1[s := A]$ , et  $g = f^{-1}$ . Par récurrence directe, *Indiv*( $s, s_1$ ), c'est-à-dire  $s$  est toujours indivisible dans  $s_1$ . Donc  $N_1 \rightarrow N_2$  si et seulement si  $f(N_1) \rightarrow f(N_2)$ . Les conditions du Lem. 4 sont vérifiées, et  $N \sim_{\Omega, \mathcal{I}} N'$ .

**(ESPÈCE DÉPENDANTE)** : On va appliquer le Lem. 5.

Soit  $N_1 = R \cup R'' \mid s_1 \in \{(N \mid N'') \rightarrow^*\}$  un réseau. Remarquons premièrement que par récurrence directe, on a  $s_1(A) \geq s_1(B)$ . On définit la fonction  $f(N_1) = R \setminus A \cup R'' \mid s_1 \setminus A$ . Si  $N_1 \rightarrow^* N_2$ , alors trivialement  $f(N_1) \rightarrow^* f(N_2)$  (condition 3). Réciproquement, pour tout réseau  $N'_1 = R \setminus A \cup R'' \mid s'_1 \in \{(N' \mid N'') \rightarrow^*\}$ , on définit l'ensemble des réseaux  $g(N'_1) = \{R \cup R'' \mid s'_1 + a'_1 A \mid a'_1 \geq s'_1(B)\}$ . Si  $N'_1 \rightarrow_{r'} N'_2$ , alors pour tout réseau  $N_1 \in g(N'_1)$ , il existe  $N_2 \in g(N'_2)$  tel que  $N_1 \rightarrow_r N_2$ . En effet, si  $r' \in R''$ , alors  $r = r'$ . Si  $r' \in R \setminus A$ , alors il existe  $r = s_1 \rightarrow s_2 \in R$  telle que  $r \setminus A = r'$ . Or comme la réaction  $r'$  était applicable, on a  $s'_1(B) \geq s_1(B)$ . D'après les conditions de la règle,  $s_1(B) \geq s_1(A)$ , et donc  $s(N_1)(A) \geq s'_1(B) \geq s_1(B) \geq s_1(A)$ , donc il y a suffisamment de molécules de  $A$ , et la réaction  $r$  est bien applicable. On obtient la solution  $s(N_2)(A) = s(N_1)(A) - s_1(A) + s_2(A) \geq s'_1(B) - s_1(B) + s_2(B) = s(N'_2)(B)$ , donc  $N_2 \in g(N'_2)$ . Par récurrence immédiate, on vérifie la condition 4 du lemme.

Notons qu'on a aussi  $N_1 \in g(f(N_1))$ , et  $f(g(N'_1)) = \{N'_1\}$  (conditions 5 et 6). Ces fonctions préservent l'observation :  $\Omega(f(N_1)) = \Omega(N_1)$ , et pour tout  $N_1 \in g(N'_1)$ ,  $\Omega(N_1) = \Omega(N'_1)$  (condition 7). Enfin,  $N \in g(N')$  et  $N' = f(N)$  (conditions 1 et 2). Donc on peut appliquer le Lem. 5, et  $N \sim_{\Omega, \mathcal{I}} N'$ .  $\square$

**Décidabilité et complexité** On étudie brièvement la décidabilité et la complexité des différentes règles. La plupart des règles de simplification sont trivialement décidables, et ne demandent qu'à parcourir l'ensemble des réactions en testant

les conditions de la règle. Elles sont donc en générale linéaire ou quadratique (lorsqu'on combine des réactions, par exemple dans (INTERMÉDIAIRE - RÉACTANT)).

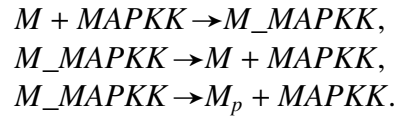
Dans le cas des règles (RÉDUCTION INVERSIBLE), (DÉCOMPOSITION - SUPPRESSION) et (DÉCOMPOSITION - SIMPLIFICATION), il faut également calculer une partie de la dynamique du réseau, afin de vérifier les conditions de type  $(R \mid s) \rightarrow^*(R \mid s')$ . Ces conditions peuvent se ramener au problème d'accessibilité dans les réseaux de Pétri, qui est décidable (Mayr, 1984; Kosaraju, 1982), mais est EXPSPACE-dur (Lipton, 1976). De plus, tous les algorithmes connus sont non-primitif récursif (Mayr, 1984; Kosaraju, 1982; Esparza, 1998). Dans le cas de (DÉCOMPOSITION - SIMPLIFICATION), il faut faire ces tests d'accessibilités pour toutes les solutions  $s'_1$  et  $s'_2$  telles que  $|s'_1 + s'_2| < |s_1 + s_2|$  (qui sont donc en nombre fini). Pour la règle (RÉDUCTION INVERSIBLE), pour une espèce  $A$ , il faut trouver une solution  $s$  (de même observation) telle que  $(R \mid A) \rightarrow^*(R \mid s)$  et réciproquement. Cela revient à trouver un marquage accessible et co-accessible depuis le marquage initial dans un réseau de Pétri. L'ensemble de ces marquages forme un ensemble semi-linéaire (Araki & Kasami, 1977), et donc savoir s'il est vide ou non est décidable. S'il est non vide, on peut alors trouver une telle solution  $s$  en testant toutes les solutions appartenant à  $\{(R \mid A) \rightarrow^k\}$  et en incrémentant progressivement  $k$ . Donc cette règle est décidable.

De façon plus générale, toutes les notions d'équivalences classiques (bisimulation, équivalence de traces, etc.) sur les réseaux de Pétri sont indécidables (Esparza, 1998). Il est donc probable que l'équivalence et la congruence des attracteurs soient également des problèmes indécidables.

### 3.4 Comparaison de modèles : l'exemple de MAPK

On va maintenant utiliser les règles de simplification afin de comparer différents modèles d'un même réseau. La méthode consiste à simplifier les réseaux, puis à comparer directement les réseaux simplifiés entre eux. On illustre cette comparaison sur 6 réseaux du système MAPK.

Ces réseaux sont les modèles *BIOMD0000000026* à *BIOMD0000000031* (que l'on notera juste 26 à 31 par commodité) de la base de données de modèles biologiques [Biomodels](#). Ils sont tous originaires de l'article ([Markevich et al., 2004](#)). Ils représentent tous la double phosphorylation de deux sites d'une espèce  $M$ , formant ainsi l'espèce  $M_{pp}$ . La phosphorylation se fait en présence d'un activateur  $MAPKK$ , tandis qu'une déphosphorylation est possible en présence de l'activateur  $MKP$ . Selon les modèles, on peut choisir de distinguer quel site est phosphorylé (on a ainsi les espèces  $M_{pY}$  et  $M_{pT}$ , comme dans le réseau de l'Ex. 9), ou de considérer juste que l'espèce est une fois phosphorylé, sans savoir quel site (on a alors juste l'espèce  $M_p$ ). De plus, les étapes de phosphorylation et déphosphorylation peuvent soit être faites de façon directe, comme dans l'Ex. 9, ou bien à l'aide d'une chaîne de réactions similaires au cas de Michaelis-Menten, par exemple :



La Fig. 3.12 indique quel modèle se trouve dans quelle catégorie. Ainsi par exemple, les réseaux 26, 27 et 31 ont une seule espèce  $M_p$ , tandis que les autres réseaux ont deux espèces  $M_{pY}$  et  $M_{pT}$ . Cette figure indique également le nombre d'espèces et de réactions du réseau initialement. On rappelle que le réseau présenté dans l'Ex. 3.12 est le modèle 29.

Le réseau 31 se distingue du réseau 27 (et des autres) par la présence de deux activateurs différents  $MAPKK_1$  et  $MAPKK_2$ , le premier étant utilisé dans la phosphorylation de  $M$  en  $M_p$ , et le second pour celle de  $M_p$  en  $M_{pp}$ . Tous les autres réseaux n'utilisent qu'un seul activateur  $MAPKK$  pour ces deux étapes. Dans le réseau 28, contrairement aux réseaux 39 et 30, l'espèce  $M_{pp}$  ne peut se déphosphoryler qu'en  $M_{pT}$ , et non pas en  $M_{pY}$ .

La fonction d'observation étend celle définie dans l'Ex. 12. Elle ne peut pas distinguer quel site est phosphorylé, c'est-à-dire :

$$\Omega_{MAPK}(M_p) = \Omega_{MAPK}(M_{pT}) = \Omega_{MAPK}(M_{pY}) = M_p.$$

Elle ne distingue pas les différents activateurs du réseau 31, c'est-à-dire :

$$\Omega_{MAPK}(MAPKK_1) = \Omega_{MAPK}(MAPKK_2) = \Omega_{MAPK}(MAPKK) = MAPKK.$$

	1 site : $M_p$	2 sites : $M_{pY}$ et $M_{pT}$
Réactions directes	27 ( $n = 5, m = 4$ ) 31 ( $n = 6, m = 4$ )	29 ( $n = 6, m = 7$ )
Réactions indirectes	26 ( $n = 11, m = 16$ )	28 ( $n = 16, m = 27$ ) 30 ( $n = 18, m = 32$ )

Fig 3.12: Classification des réseaux du système MAPK, avec  $n$  le nombre d'espèce et  $m$  celui de réactions.

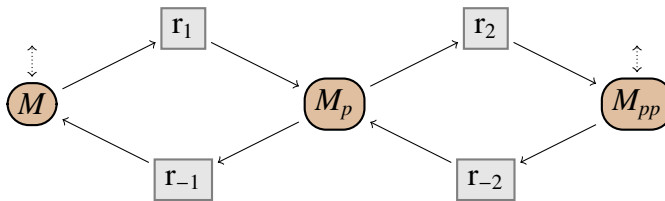


Fig 3.13: Graphe du réseau simplifié de MAPK.

Enfin, l'observation ne voit pas si les espèces sont libres ou dans un complexe. Par exemple :

$$\Omega_{MAPK}(M\_MAPKK) = \Omega_{MAPK}(M) + \Omega_{MAPK}(MAPKK) = M + MAPKK.$$

L'observation est étendue aux réseaux de façon naturelle :

$$\Omega_{MAPK}(s) = \sum_{A \in s} s(A) \Omega_{MAPK}(A).$$

On considère de plus que seules les espèces  $M$  et  $M_{pp}$  peuvent interagir avec le contexte, et donc que toutes les autres espèces sont internes. Enfin, tous les réseaux possèdent initialement des molécules de  $M$ , et des molécules des activateurs  $MAPKK$  (ou de  $MAPKK_1$  et  $MAPKK_2$  pour le réseau 31) et  $MKP$ .

On peut simplifier ces réseaux en utilisant principalement deux types de règles. D'une part, les règles de symétrie (**SYMÉTRIE - RÉACTANTS**) et (**SYMÉTRIE - PRODUITS**), afin de fusionner les espèces  $M_{pY}$  et  $M_{pT}$ . Par exemple, le modèle 28 peut être simplifier pour obtenir le modèle 29, représenté dans la Fig. 2.10 de l'Ex. 9. D'autre part, on peut également utiliser la règle (**RÉDUCTION INVERSIBLE**), pour simplifier les chaînes de type Michaelis-Menten en une unique réaction. Les règles (**RÉACTION CYCLIQUE**) et (**ACTIVATEUR PRÉSENT**) sont également utilisées.

En appliquant la simplification à ces six réseaux, on obtient un unique réseau simplifié, avec 4 réactions et 5 espèces. Ses réactions sont illustrées dans la Fig. 3.13. Notons que seules 3 espèces sont utilisées dans les réactions :  $M$ ,

$M_p$  et  $M_{pp}$ . Les activateurs *MAPKK* et *MKP* ne sont plus dans les réactions, mais apparaissent toujours dans la solution initiale du réseau réduit. On pourrait les supprimer avec la règle (ESPÈCE INUTILE) si elles n'étaient pas observables.

## 3.5 Discussion

Dans cette section, on discute et compare notre simplification avec d'autres simplifications dans la littérature. On étudie tout d'abord l'effet sur la congruence d'un changement de fonction d'observation. On étudie les liens entre notre équivalence et la bisimulation, ainsi qu'avec les épimorphismes de sous-graphe. Enfin, on compare notre méthode de simplification avec d'autres méthodes préservant d'autres sémantiques non déterministes, dans les réseaux de Pétri et les réseaux booléens.

### 3.5.1 Changement de fonction d'observation

On étudie les caractéristiques de la congruence  $\sim_{\Omega, I}$  selon l'observation  $\Omega$ .

**Observation totale ou aveugle** On considère tout d'abord deux observation particulières, l'observation *totale*, qui voit tout, et l'observation *aveugle*, qui ne voit rien.

L'observation totale,  $\Omega_{tot}$ , est définie pour tout réseau  $N$  par  $\Omega_{tot}(N) = s(N)$ . Avec cette fonction d'observation, un ensemble de réseaux  $\mathcal{A}$  est un attracteur s'il est fermé et connecté, c'est-à-dire pour tout  $N \in \mathcal{A}$ , on a exactement  $\mathcal{A} = \{N \rightarrow^*\}$ . Dans la Fig. 3.1, avec cette fonction d'observation, les réseaux  $N_1, N_4, N_5$  et  $N_6$  ne sont plus équivalents. Les réseaux  $N_1, N_2$  et  $N_3$  sont encore équivalents, puisque seules la partie intermédiaire diffère. Les réseaux divergeant  $N_7, N_8, N_9$  et  $N_{10}$  sont tous équivalents.

À l'inverse, l'observation aveugle  $\Omega_{av}$  ne peut rien observer, c'est-à-dire pour tout réseau  $N$ ,  $\Omega_{av}(N) = \mathbb{O}$ . Dans ce cas, la propriété de  $\Omega$ -connectivité est toujours vérifiée, donc tout ensemble  $\mathcal{A}$  fermé est un attracteur. En particulier, cela implique qu'on ne peut jamais diverger. Et tous les attracteurs ont la même observation  $\{\mathbb{O}\}$ . Donc pour tout réseau  $N$ , on a  $N \downarrow_{\{\mathbb{O}\}}^{\Omega_{av}}$  et  $N \uparrow^{\Omega_{av}}$ . Donc tous les réseaux sont attracteurs-congruents pour l'observation aveugle.

**Hiérarchie entre les observations** De façon plus générale, on peut définir une hiérarchie entre les fonctions d'observation. Plus une fonction d'observation est petite, moins elle va pouvoir distinguer des réseaux :

$$\Omega_1 \leq \Omega_2 \quad \text{si et seulement si} \quad (\Omega_2(N) = \Omega_2(M) \text{ implique } \Omega_1(N) = \Omega_1(M)).$$

Pour toute fonction d'observation  $\Omega$ , on a donc  $\Omega_{av} \leq \Omega \leq \Omega_{tot}$ .

**Exemple 30.** Dans l'Ex. 11, on a  $\Omega'_{Tet-On} \leq \Omega_{Tet-On}$ . Rappelons que ces fonctions sont définies par :

$$\Omega_{Tet-On}(nGFP_a + s) = nGFP_a \quad \text{où } GFP_a \notin s,$$

$$\Omega'_{\text{Tet-On}}(nGFP_a + s) = \min(n, 1) \quad \text{où } GFP_a \notin s.$$

En effet, si le nombre de molécules de  $GFP_a$  est identique dans deux réseaux, alors la présence ou l'absence de  $GFP_a$  dans ces réseaux est également la même.

On peut alors remarquer que les attracteurs et les convergences sont préservés lorsque la fonction d'observation décroît, tandis que la divergence est conservée lorsque l'observation croît.

**Proposition 6.** Si  $\Omega_1 \leq \Omega_2$ , et que  $\mathcal{A}$  est un  $\Omega_2$ -attracteur, alors  $\mathcal{A}$  est aussi un  $\Omega_1$ -attracteur.

*Proof.* Soit  $\mathcal{A}$  un  $\Omega_2$ -attracteur. Notons tout d'abord que la propriété de clôture est préservée quelque soit l'observation. Soit  $N \in \mathcal{A}$  et  $o_1 \in \Omega_1(\mathcal{A})$  une observation. Il existe donc un réseau  $N' \in \mathcal{A}$  tel que  $o_1 = \Omega_1(N')$ . On définit  $o_2 = \Omega_2(N')$ . Par la propriété de  $\Omega_2$ -connectivité, on a  $o_2 \in \Omega_2(\{N \rightarrow^*\})$ . Il existe donc un réseau  $N'' \in \{N \rightarrow^*\}$  tel que  $\Omega_2(N'') = o_2 = \Omega_2(N')$ . Mais comme  $\Omega_1 \leq \Omega_2$ , cela implique  $\Omega_1(N'') = \Omega_1(N') = o_1$ . Donc  $o_1 \in \Omega_1(\{N \rightarrow^*\})$ , et  $\mathcal{A}$  est un  $\Omega_1$ -attracteur.  $\square$

**Corollaire 1.** Soit  $\Omega_1 \leq \Omega_2$ . S'il existe un attracteur  $\mathcal{A}$  tels que  $N \downarrow_{\Omega_2(\mathcal{A})}^{\Omega_2}$ , alors  $N \downarrow_{\Omega_1(\mathcal{A})}^{\Omega_1}$ .

Enfin, remarquons que l'ordre est inversée pour les divergences.

**Proposition 7.** Si  $\Omega_1 \leq \Omega_2$ , et que  $N \uparrow^{\Omega_1}$ , alors  $N \uparrow^{\Omega_2}$ .

*Proof.* Supposons que  $N \uparrow^{\Omega_1}$ , c'est-à-dire qu'il existe  $N' \in \{N \rightarrow^*\}$  tel que pour tout  $O$ , on ait  $N' \not\downarrow_O^{\Omega_1}$ . Alors par la contraposition du corollaire 1 impose  $N' \not\downarrow_O^{\Omega_2}$ , et donc  $N \uparrow^{\Omega_2}$ .  $\square$

En revanche, les autres sens ne sont pas vérifiés. Par exemple, considérons les observations  $\Omega_{av} \leq \Omega_{tot}$ , et le réseau  $N = \{\emptyset \rightarrow A\} \mid \emptyset$ . On a  $N \downarrow_{\{\emptyset\}}^{\Omega_{av}}$  tandis que  $N \not\downarrow_{\{\emptyset\}}^{\Omega_{tot}}$ , et de même  $N \uparrow^{\Omega_{tot}}$  tandis que  $N \not\uparrow^{\Omega_{av}}$ . Il n'y a donc pas de lien entre l'ordre des fonctions d'observation et une inclusion des congruences des attracteurs.

**Observations équivalentes** Il est par contre possible de définir une équivalence entre les fonctions d'observation :  $\Omega_1 \sim \Omega_2$  si  $\Omega_1 \leq \Omega_2$  et  $\Omega_2 \leq \Omega_1$ . En d'autres termes, ces fonctions d'observation distinguent les mêmes réseaux : pour toute paire  $N$  et  $N'$ , on a  $\Omega_1(N) = \Omega_1(N')$  si et seulement si  $\Omega_2(N) = \Omega_2(N')$ . Par exemple, la fonction d'observation  $\Omega_{\text{Tet-On}}$  de l'Ex. 11, telle que :

$$\Omega_{\text{Tet-On}}(nGFP_a + s) = nGFP_a \quad \text{où } GFP_a \notin s,$$

est équivalente à la fonction  $\Omega''_{\text{Tet-On}}$ , définie comme en morphisme des solutions vers le monoïde des entiers naturels par :

$$\Omega''_{\text{Tet-On}}(nGFP_a + s) = n \quad \text{où } GFP_a \notin s.$$

Les relations de congruence qui découlent d'observations équivalentes sont alors identiques.

**Proposition 8.** Si  $\Omega_1 \sim \Omega_2$ , alors pour toute paire  $N$  et  $N'$ , on a  $N \sim_{\Omega_1, \mathcal{I}} N'$  si et seulement si  $N \sim_{\Omega_2, \mathcal{I}} N'$ .

*Proof.* Soit  $N \sim_{\Omega_1, \mathcal{I}} N'$ . Par la Prop. 7, on a  $N \uparrow^{\Omega_2}$  si et seulement si  $N' \uparrow^{\Omega_2}$ . De plus, pour tout  $O_2$ ,  $N \downarrow_{O_2}^{\Omega_2}$  si et seulement si il existe un  $\Omega_2$ -attracteur  $\mathcal{A} \subseteq \{N \rightarrow^*\}$  tel que  $O_2 = \Omega_2(\mathcal{A})$ . Mais par la Prop. 6,  $\mathcal{A}$  est aussi un  $\Omega_1$ -attracteur, donc  $N \uparrow^{\Omega_2}$  si et seulement si  $N \downarrow_{O_1}^{\Omega_1}$ , avec  $O_1 = \Omega_1(\mathcal{A})$ . Comme  $N \sim_{\Omega_1, \mathcal{I}} N'$ , on a aussi  $N' \downarrow_{O_1}^{\Omega_1}$ , et avec le raisonnement inverse,  $N' \downarrow_{O_2}^{\Omega_2}$ . Donc  $N \sim_{\Omega_2, \mathcal{I}} N'$ .  $\square$

On peut également observer que la réciproque est vraie.

**Proposition 9.** Si pour tout  $N, N' \in \text{Nets}$ , on a  $N \sim_{\Omega_1} N'$  si et seulement si  $N \sim_{\Omega_2} N'$ , alors  $\Omega_1 \sim \Omega_2$ .

*Proof.* Soit  $N = R \mid s$  et  $N' = R' \mid s'$  deux réseaux tels que  $\Omega_1(N) = \Omega_1(N')$ , c'est-à-dire  $\Omega_1(s) = \Omega_1(s') = o_1$ . On considère les réseaux sans réactions  $N_1 = \emptyset \mid s$  et  $N'_1 = \emptyset \mid s'$ . On a trivialement  $N_1 \downarrow_{\{o_1\}}^{\Omega_1}$ , et pour tout  $O_1 \neq \{o_1\}$ ,  $N_1 \not\downarrow_{O_1}^{\Omega_1}$ . De plus,  $N_1 \not\uparrow^{\Omega_1}$ . Les mêmes prédicats sont vérifiés pour  $N'_1$ . Donc  $N_1 \sim_{\Omega_1, \mathcal{I}} N'_1$ , et par conséquent  $N_1 \sim_{\Omega_2, \mathcal{I}} N'_1$ . On a alors  $N_1 \downarrow_{\{\Omega(2)(s)\}}^{\Omega_2}$  et  $N'_1 \downarrow_{\{\Omega(2)(s')\}}^{\Omega_2}$ . Comme ce sont les seuls ensembles d'observation vers lesquels ces réseaux convergent, cela implique  $\Omega(2)(s) = \Omega(2)(s')$ , et donc  $\Omega_2(N) = \Omega_2(N')$ . Donc pour toute paire de réseaux  $N, N' \in \text{Nets}$ ,  $\Omega_1(N) = \Omega_1(N')$  implique  $\Omega_2(N) = \Omega_2(N')$ , donc  $\Omega_2 \leq \Omega_1$ . La réciproque est symétrique, et donc  $\Omega_1 \sim \Omega_2$ .  $\square$

### 3.5.2 Bisimulation

On étudie ici le lien entre l'équivalence des attracteurs et la bisimulation (faible), une notion d'équivalence classiques dans les systèmes concurrents. On étudie également l'effet de la simplification sur la bisimulation.

On définit la relation  $\xrightarrow{\varrho}$ , qui représente la sémantique opérationnelle non observable, c'est-à-dire l'évolution possible du réseau sans que l'on puisse observer de changement :

$$\frac{N \rightarrow N' \quad \Omega(N) = \Omega(N')}{N \xrightarrow{\varrho} N'}$$



**Définition 14.** Une relation  $\mathcal{R} : \{N \rightarrow^* \} \times \{N' \rightarrow^* \}$  est une  $\Omega$ -bisimulation entre deux réseaux  $N$  et  $N'$  si :

- $N\mathcal{R}N'$ ,
- $\mathcal{R}$  préserve l'observation, c'est-à-dire que pour toute paire  $N_1\mathcal{R}N'_1$ ,  $\Omega(N_1) = \Omega(N'_1)$ ,
- pour toute paire  $N_1\mathcal{R}N'_1$ , si  $N_1 \rightarrow N_2$  avec  $\Omega(N_1) \neq \Omega(N_2)$ , alors il existe  $N'_2$  tel que  $N'_1 \xrightarrow[\varnothing]^* \rightarrow \xrightarrow[\varnothing]^* N'_2$  et  $N_2\mathcal{R}N'_2$ ,
- réciproquement, si  $N_1\mathcal{R}N'_1$  et  $N'_1 \rightarrow N'_2$  avec  $\Omega(N'_1) \neq \Omega(N'_2)$ , alors il existe  $N_2$  tel que  $N_1 \xrightarrow[\varnothing]^* \rightarrow \xrightarrow[\varnothing]^* N_2$  et  $N_2\mathcal{R}N'_2$ ,
- pour toute paire  $N_1\mathcal{R}N'_1$ , si  $N_1 \xrightarrow[\varnothing]^* N_2$ , alors il existe  $N'_2$  tel que  $N'_1 \xrightarrow[\varnothing]^* N'_2$  et  $N_2\mathcal{R}N'_2$ ,
- réciproquement, si  $N_1\mathcal{R}N'_1$  et  $N'_1 \xrightarrow[\varnothing]^* N'_2$ , alors il existe  $N_2$  tel que  $N_1 \xrightarrow[\varnothing]^* N_2$  et  $N_2\mathcal{R}N'_2$ .

Deux réseaux  $N$  et  $N'$  sont  $(\Omega, \mathcal{I})$ -bisimilaires, noté  $N \sim_{\Omega, \mathcal{I}}^{bis} N'$ , si dans tout contexte  $N''$  tel que  $Espèces(N'') \cap \mathcal{I} = \emptyset$ , il existe une  $\Omega$ -bisimulation entre  $N \mid N''$  et  $N' \mid N''$ .

La bisimulation implique la congruence des attracteurs. En revanche, pour une fonction d'observation non aveugle, deux réseaux peuvent être attracteur-équivalents sans être bisimilaires. Pour la fonction d'observation aveugle  $\Omega_{av}$ , tous les réseaux sont trivialement bisimilaires et attracteurs-congruents, donc on a  $\sim_{\Omega_{av}, \mathcal{I}}^{bis} = \sim_{\Omega_{av}, \mathcal{I}}$ .

**Proposition 10.** Pour tout  $\Omega \neq \Omega_{av}$ , on a  $\sim_{\Omega, \mathcal{I}}^{bis} \subsetneq \sim_{\Omega, \mathcal{I}}$ .

*Proof.* Pour  $\sim_{\Omega, \mathcal{I}}^{bis} \subset \sim_{\Omega, \mathcal{I}}$ , la preuve est une variante du Lem. 5, dans laquelle la fonction  $f$  produit aussi un ensemble de réseaux :  $f : \{N \rightarrow^* \} \mapsto 2^{\{N' \rightarrow^* \}}$ . Avec  $\mathcal{R}$  une bisimulation, on prend alors  $f(N_1) = \{N'_1 \mid N_1\mathcal{R}N'_1\}$ , et réciproquement pour  $g$ . Les conditions du lemme sont alors trivialement vérifiées par définition de la bisimulation.

Pour  $\sim_{\Omega, \mathcal{I}}^{bis} \neq \sim_{\Omega, \mathcal{I}}$ , on considère les réseaux  $N = \{A \rightarrow B, B \rightarrow C, B \rightarrow D\} \mid A$  et  $N' = \{A \rightarrow B_1, B_1 \rightarrow C, A \rightarrow B_2, B_2 \rightarrow D\} \mid A$ , illustrés dans la Fig. 3.14 avec  $\mathcal{I} = \{A, B, B_1, B_2, C, D\}$  et  $\Omega(B) = \Omega(B_1) = \Omega(B_2)$ . C'est un exemple classique de non-bisimilarité :  $N \not\sim_{\Omega, \mathcal{I}}^{bis} N'$ . En revanche, les deux réseaux sont bien attracteurs-congruents : ils peuvent tous les deux converger vers  $\{\Omega(C)\}$ ,  $\{\Omega(D)\}$  et ne peuvent

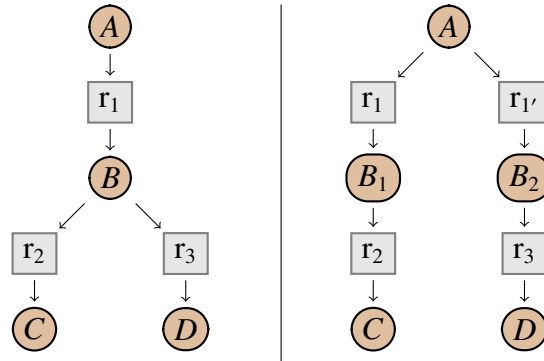


Fig 3.14: Exemple de réseaux attracteur-équivalents et non bisimilaires.

pas diverger. On peut aussi considérer l'exemple  $N = \{A \rightarrow B\} \mid A$  et  $N' = \emptyset \mid B$ , qui ont le même comportement final (donc sont attracteurs-congruents) mais pas le même comportement transitoire (donc pas bisimilaires).  $\square$

Certaines règles de simplification sont correctes pour la bisimulation. C'est le cas directement des règles (**RÉACTION CYCLIQUE**), (**ACTIVATEUR PRÉSENT**), (**ESPÈCE INUTILE**), (**RÉACTIONS INAPPLICABLES**), (**PRODUCTIONS GROUPÉES**), (**SOLUTION INDIVISIBLE**), (**ESPÈCE DÉPENDANTE**). C'est également le cas pour (**RÉACTION INVERSIBLE**), comme les réactions inversibles ne sont pas observables, mais pas pour (**RÉDUCTION INVERSIBLE**), puisque la séquence  $(R(N) \mid A) \xrightarrow{*} (R(N) \mid s)$  (ou sa réciproque) peut être observable. Si on remplaçait cette condition par  $R(N) \mid A \xrightarrow{*} R(N) \mid s$ , la règle deviendrait correcte pour la bisimulation. En ajoutant des conditions similaires sur l'observation, les règles de décomposition, ainsi que la règle (**SYMÉTRIE - RÉACTANTS**), pourraient également devenir correctes pour la bisimulation. En revanche, elle n'est pas préservée par (**CASCADE**), ni par les règles d'élimination des espèces intermédiaires. Par exemple, le réseau  $\{A_1 \rightarrow A_2, A_2 \rightarrow B, A_1 \rightarrow C\} \mid A_1$  se simplifie avec ces règles en  $\{A_1 \rightarrow B, A_1 \rightarrow C\} \mid A_1$ . Or ces réseaux ne sont pas bisimilaires, puisque dans le premier il existe un réseau avec pour solution  $A_2$ , où l'on ne pourra que produire du  $B$  et pas de  $C$ . Une telle solution n'existe pas dans le deuxième réseau. Pour (**SYMÉTRIE - PRODUITS**), l'exemple classique de non-bisimulation dans la preuve ci-dessus montre que ce n'est pas non plus préservée. De même, elle n'est pas préservée par (**SYMÉTRIE - IMITATION**), puisque la solution  $s$  peut avoir des comportements supplémentaires par rapport à l'espèce  $A$ , et donc  $A$  et  $s$  ne peuvent pas être bisimilaires.

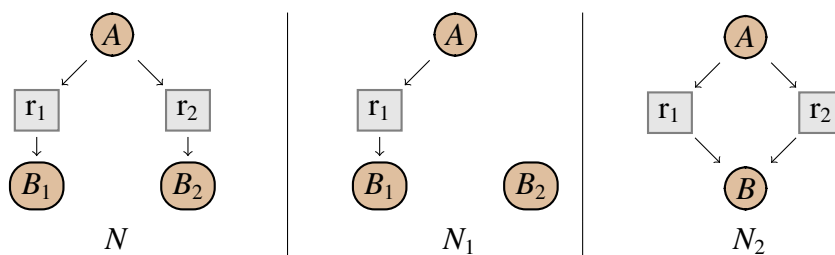


Fig 3.15: Réseau de réactions  $N$ . Le réseau  $N_1$  est obtenu en éliminant la réaction  $r_2$ , et le réseau  $N_2$  en fusionnant les espèces  $B_1$  et  $B_2$ .

### 3.5.3 Épimorphisme de sous-graphe

Dans (Gay et al., 2010), les auteurs proposent une méthode de simplification de réseaux biologiques basée sur l'élimination et la fusion d'espèces et de réactions. Par exemple, considérons le réseau  $N$  représenté graphiquement dans la Fig. 3.15. On peut le simplifier en éliminant la réaction  $r_2$ , et obtenir le réseau  $N_1$ , ou en fusionnant les espèces  $B_1$  et  $B_2$ , et obtenir le réseau  $N_2$ .

Avec ces opérations, un réseau  $N'$  est la simplification d'un réseau  $N$  si et seulement s'il existe un épimorphisme de sous-graphe entre  $N$  et  $N'$  (Gay et al., 2014). Un épimorphisme de  $N$  vers  $N'$  est un morphisme surjectif des ensembles des espèces et des réactions de  $N$  vers ceux de  $N'$ . Un épimorphisme de sous-graphe est un morphisme surjectif allant d'un sous-ensemble des espèces et des réactions de  $N$  vers l'ensemble des espèces et des réactions de  $N'$ . Par exemple, les fonctions  $f_1$  et  $f_2$  définies ci-dessous sont des épimorphismes de sous-graphe entre les réseaux  $N$  et respectivement  $N_1$  et  $N_2$  de la Fig. 3.15.

$$\left\{ \begin{array}{l} f_1(A) = A \\ f_1(B_1) = B_1 \\ f_1(B_2) = B_2 \\ f_1(r_1) = r_1 \end{array} \right. \quad \left\{ \begin{array}{l} f_2(A) = A \\ f_2(B_1) = B \\ f_2(B_2) = B \\ f_2(r_1) = r_1 \\ f_2(r_2) = r_2 \end{array} \right.$$

Plusieurs de nos règles de simplification sont également basées sur ces opérations d'élimination ou de fusion d'espèces et de réactions. Il est néanmoins nécessaire d'ajouter des conditions supplémentaires, afin que l'équivalence des attracteurs soit préservée. Par exemple, on peut supprimer une espèce que si celle-ci n'est pas utile, avec la règle (ESPÈCE INUTILE). De même, on ne peut fusionner deux espèces que si elles présentent certaines propriétés de symétrie. Sous certaines conditions, les règles (RÉACTION CYCLIQUE), (DÉCOMPOSITION - SUPPRESSION) ou (RÉACTIONS INAPPLICABLES) permettent de supprimer une réaction.

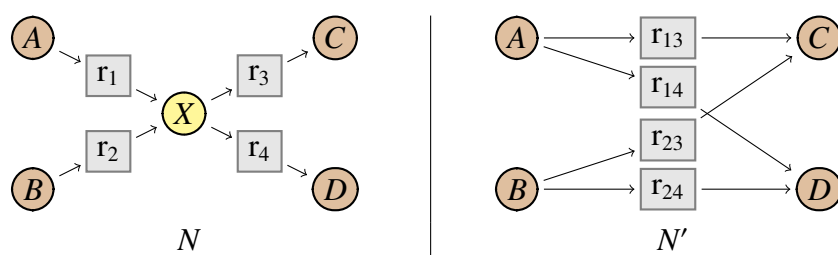


Fig 3.16: Simplification du réseau  $N$  en  $N'$  avec la règle (INTERMÉDIAIRE - RÉACTANT).

À l'inverse, certaines de nos règles de simplification ne peuvent être obtenues en combinant des éliminations et des fusions d'espèces et de réactions. Donc dans ce cas il n'existe pas d'épimorphisme de sous-graphe entre le réseau initial et le réseau simplifié. C'est par exemple le cas de certaines applications des règles d'élimination d'espèces intermédiaires. Il faut en effet pouvoir fusionner des ensembles de réactions entre eux, et pas simplement une réaction avec une autre. Par exemple, on peut simplifier le réseau  $N$  (Fig. 3.16) avec la règle (INTERMÉDIAIRE - RÉACTANT), et obtenir le réseau  $N'$ , en fusionnant deux-à-deux les réactions de  $\{r_1, r_2\}$  avec celles de  $\{r_3, r_4\}$ .

### 3.5.4 Réseaux de Pétri

Comme on l'a vu, il existe un lien direct entre les réseaux de réactions biologique et les réseaux de Pétri (Reddy et al., 1993; Chaouiya, 2007; Heiner et al., 2008). Il est donc naturel de comparer notre simplification avec ce qui existe dans la littérature des réseaux de Pétri. De nombreuses approches ont été proposées pour simplifier les réseaux de Pétri, par exemple utilisant des propriétés de bisimilarité (Autant & Schnoebelen, 1992; Schnoebelen & Sidorova, 2000), ou des règles de réécriture proches des nôtres (Berthelot, 1985; Lee-Kwang et al., 1987).

En général, ces simplifications visent à préserver certaines propriétés spécifiques des réseaux de Pétri, telles que la vivacité, le caractère borné, ou encore des propriétés d'accessibilité. On dit qu'un réseau est *vivace* si pour toute réaction (transition) et à partir de tout réseau accessible, on peut atteindre un autre réseau où la réaction est applicable. Un réseau est *borné* si le nombre total de molécules (jetons) dans les réseaux accessibles est borné par une constante. Un réseau est *sans-impasse* si dans n'importe quel réseau accessible, on peut toujours appliquer une réaction et atteindre un autre réseau. Un réseau  $N$  possède un *état d'accueil*  $N'$  si  $N'$  est accessible depuis tous les réseaux de  $\{N \rightarrow^*\}$ . Un réseau  $N$  est *réversible* si  $N$  est lui-même un état d'accueil. Remarquons qu'un réseau est réversible si et

seulement si  $\{N \rightarrow^*\}$  est un  $\Omega_{tot}$ -attracteur.

Même si les motivations sont différentes, on peut remarquer qu'on retrouve le même type de simplification pour ces réseaux. Par exemple, Berthelot et al. a développé un ensemble de règles de simplification préservant ces propriétés (Berthelot & Roucairol, 1976; Berthelot et al., 1980; Berthelot, 1985). Ses règles, comme celles présentées dans cette thèse, sont basées sur l'élimination de places (espèces) intermédiaires, de dépendances ou de symétries. Par exemple, la règle de substitution (R1) dans (Berthelot & Roucairol, 1976) vise à fusionner un ensemble de réactions produisant une solution  $s$  avec un ensemble de réactions consommant  $s$ , lorsque les espèces de  $s$  ne sont jamais utilisées ailleurs. Elle est équivalente (modulo les conditions de contexte et d'observation spécifiques à notre sémantique) à transformer la solution  $s$  en une espèce  $A$  avec la règle (SOLUTION INDIVISIBLE), puis à fusionner les réactions avec la règle (INTERMÉDIAIRE - RÉACTANT). Le règle d'élimination des places redondantes (R2) correspond à peu près à la règle (ESPÈCE DÉPENDANTE), tandis que la règle d'élimination des transitions sans importance (R3) correspond à la règle (RÉACTION CYCLIQUE). Certaines règles traitent également le cas des symétries, comme par exemple la règle (TP3) dans (Berthelot, 1985), similaire à la règle (SYMÉTRIE - RÉACTANTS). Notons que les règles présentent néanmoins à chaque fois de légères différences. Ainsi, nos règles demandent en générale des conditions supplémentaires sur l'observation ou les espèces internes. À l'inverse, la règle (R3) ne supprime que les transitions (réactions) cycliques dont on sait qu'elles seront toujours applicables plus tard, afin de ne pas rendre vivace un réseau qui ne l'était pas. Ainsi, bien que proches, nos règles ne préservent pas en général les propriétés usuelles des réseaux de Pétri, et réciproquement les règles de Berthelot et al. ne préservent pas la congruence des attracteurs.

Lee et al. (Lee & Favrel, 1985; Lee-Kwang et al., 1987) ont également proposé un ensemble de règles de simplifications préservant les propriétés classiques des réseaux de Pétri. Ces règles sont basées encore une fois sur l'élimination des espèces intermédiaires, les symétries, ou la fusion de cycles (comme on peut le faire avec la règle (RÉDUCTION INVERSIBLE)). Le même genre de simplification permet également de préserver certaines propriétés d'équité (voir par exemple (Silva & Murata, 1992)). Les travaux de Berthelot et al. ont été généralisés pour les réseaux Prédicat/Transitions (Genrich, 1988; Krohn & Rauhamaa, 1988), les réseaux de Pétri colorés (Colom, Martinez, & Silva, 1986; Haddad, 1991), les réseaux de Pétri temporels (Sloan & Buy, 1996) et les réseaux Ada (Shatz, Tu, Murata, & Duri, 1996). Dans (Haddad & Pradat-Peyre, 2006), ces règles sont également généralisées afin de préserver également des propriétés de la logique LTL. Enfin, certaines règles de simplification ont été développées pour des sous-classes spécifiques de réseaux de Pétri. Dans (Murata & Koh, 1980; Johnsonbaugh & Murata, 1981), Murata et al. présentent des règles de simplification spécifiques

pour les graphes marqués fortement connexes, des réseaux de Pétri particulier dans lesquels chaque place à au maximum une transition entrante et une transition sortante. Dans (Desel, 1990; Esparza & Silva, 1990; Kovalyov, 1990) des simplifications spécifiques ont été proposées pour le cas des réseaux de Pétri bornés et avec libre choix. Des simplifications similaires ont également été proposées pour les IO-graphes (Juan, Tsai, & Murata, 1998) ou les architectures orientées service (Wolf, Stahl, Ott, & Danitz, 2009).

### 3.5.5 Réseaux booléens

Un autre formalisme couramment utilisé en modélisation biologique est celui des réseaux booléens ou multivalués, en particulier pour représenter des réseaux de gènes. Dans ces réseaux, les nœuds codent pour des gènes, et des arcs représentent l'effet d'un gène sur un autre, représenté par une fonction logique. Chaque gène peut avoir soit deux états (cas booléen), soit un nombre borné (cas multivalué), contrairement à nos espèces qui peuvent avoir n'importe quel nombre de molécules. Plusieurs sémantiques opérationnelles sont alors possibles : on peut par exemple choisir une sémantique synchrone, comme dans les réseaux de Kauffman (Kauffman, 1969, 1993). Dans ce cas, on applique toutes les réactions d'un seul coup, en mettant tous les gènes à jour en une seule étape. Cette sémantique est donc déterministe. Au contraire, dans les réseaux de Thomas (Thomas, 1973, 1991), on met un seul gène à jour à chaque étape. Il s'agit alors d'une dynamique asynchrone et non déterministe. On peut également choisir une dynamique intermédiaire, en mettant à jour un sous-ensemble de gènes à chaque étape.

Il est alors fréquent d'étudier le comportement final des réseaux, c'est-à-dire l'accessibilité des composantes fortement connexes terminales. En comparaison avec nos  $\Omega$ -attracteurs, ces composantes sont toujours de taille finie. Cela implique notamment qu'un réseau ne peut pas diverger. Il n'y a pas non plus de fonction d'observations, et donc un attracteur est un ensemble  $\mathcal{A}$  tel que pour tout réseau (booléen)  $N \in \mathcal{A}$ ,  $\{N \rightarrow^*\} = \mathcal{A}$ . En général, on appelle *état stable* un attracteur de taille 1, et *cycle* un attracteur de taille strictement supérieure à 1. L'ensemble des attracteurs accessibles dépend de la sémantique choisie. En revanche, l'ensemble des états stables accessibles est le même quelque soit la sémantique.

Plusieurs méthodes de simplification cherchent alors à préserver ce comportement asymptotique. Une simplification possible consiste par exemple à supprimer les *sorties* d'un réseau, c'est-à-dire les gènes qui ne régulent aucun autre gène, et donc dont l'état n'intervient pas dans la dynamique, comme le proposent Bilke et al. (Bilke & Sjunnesson, 2001) pour les réseaux de Kauffman. Ils proposent également de supprimer des gènes dit *gelés*, c'est-à-dire qui auront toujours la même valeur dans le comportement asymptotique, quelque soit les valeurs initiales du

réseaux. Leur simplification préserve le comportement asymptotique (unique pour un réseau initial fixé, puisque la dynamique est déterministe). Richardson (Richardson, 2005) propose, toujours dans les réseaux de Kauffman, une méthode similaire, ainsi qu'une procédure permettant de déterminer les gènes gelés. Matache (Matache & Matache, 2016) propose quant à lui de hiérarchiser les gènes selon leur "pouvoir déterminant", c'est-à-dire leur impact sur le reste du réseau, et d'utiliser cet ordre pour éliminer de façon itérative les gènes les moins importants, sans trop modifier le comportement final.

Saadatpour et al. (Saadatpour, Albert, & Albert, 2010; Saadatpour, Albert, & Reluga, 2013), dans les réseaux de Thomas, proposent d'une part d'éliminer de façon itérative les *entrées* du réseau, c'est-à-dire des gènes n'étant pas régulés pas d'autres, et donc dont l'état reste toujours le même. D'autre part ils éliminent également les nœuds *intermédiaires*  $X$ , sous la condition qu'il n'existe qu'un seul gène qui influe  $X$ , et symétriquement que  $X$  n'influence qu'un seul gène. Ils prouvent que leur simplification préserve les attracteurs, et est confluente. Naldi et al. (Naldi, Remy, Thieffry, & Chaouiya, 2011; Naldi et al., 2012) proposent également de supprimer des gènes intermédiaires dans des réseaux de Thomas multivalués, et suppriment les sorties du réseau. Ils prouvent que les attracteurs accessibles dans le réseau initial le sont toujours dans le réseau simplifié. En revanche de nouveaux cycles peuvent apparaître durant la simplification. De plus, leur simplification n'est pas confluente.

Dans (Veliz-Cuba, 2011), Veliz décrit une simplification confluente de gènes intermédiaires et des sorties d'un réseau qui préserve les états stables, quelque soit la sémantique utilisée. Dans (Veliz-Cuba, Laubenbacher, & Aguilar, 2013), une méthode de simplification de réseau ET-NON (une autre représentation de réseaux booléens) est présentée, préservant le nombre d'états stables. Une représentation étendue des réseaux booléens est utilisée dans (Zañudo & Albert, 2013) afin d'identifier et d'éliminer les gènes se stabilisant. Enfin, Paulevé (Paulevé, 2016) modélise des systèmes biologiques à l'aide de réseaux d'automates, de façon similaire à des réseaux multivalués. Ces réseaux sont alors simplifiés en cherchant à préserver une propriété d'accessibilité fixée initialement, en supprimant les états n'étant pas nécessaires pour atteindre cet objectif.

On voit ainsi qu'on retrouve dans les réseaux booléens des simplifications similaires à certaines de nos règles. En plus de l'élimination des espèces intermédiaires, l'élimination des entrées peut être faite avec la règle (**INTERMÉDIAIRE - INITIALE**), tandis que celle des sorties est réalisée par la règle (**ESPÈCE INUTILE**), modulo l'observation.

## 3.6 Conclusion

Dans ce chapitre, on a présenté une nouvelle sémantique pour les réseaux de réactions chimiques. Cette sémantique s'intéresse aux comportements finaux réguliers des réseaux, représentés par des attracteurs. On a étudié la capacité des réseaux à pouvoir converger vers ces attracteurs (modulo la fonction d'observation), ou au contraire à diverger. Des notions d'équivalence et de congruence ont été définies à partir de cette sémantique. On a ensuite présenté un ensemble de règles de simplification, exacte pour la congruence des attracteurs. Ces règles permettent de simplifier efficacement un réseau, par exemple en supprimant des espèces inutiles, en fusionnant des cascades de réactions, ou en utilisant les symétries du réseau. On a prouvé qu'un sous ensemble de ces règles formait une simplification confluente. On a montré sur l'exemple de MAPK que notre simplification pouvait être utilisée pour prouver l'équivalence entre différents réseaux d'un même système. Enfin, on a comparé notre équivalence et notre simplification avec les autres méthodes existantes dans la littérature.

L'ensemble des règles de simplification proposées ici n'est pas complet pour notre notion de congruence. Il pourrait donc être intéressant de développer de nouvelles règles, en fonction de réseaux tirés de modèles biologiques concrets. Il serait également intéressant d'appliquer notre méthode de simplification de façon semi-automatique (puisque'il faut décider pour chaque modèle quelles sont les espèces internes et quelle est la fonction d'observation) sur les réseaux de la base de données [Biomodels](#), et d'utiliser les résultats pour comparer ces modèles, comme on l'a fait pour le réseau MAPK. Enfin, il semble également possible de développer des méthodes de simplification similaires pour les langages à base de règle, comme *Kappa* ([Danos & Laneve, 2004](#)) ou *React(C)* ([John et al., 2011](#)).





## CHAPITRE 4

---

### Simplification préservant la sémantique déterministe

---

On s'intéresse maintenant à la sémantique déterministe des réseaux de réactions cinétiques, telle que définie dans la Sec. 2.2. Une première étape consiste à définir formellement une notion de *similarité*, qui va nous permettre de réécrire des expressions, par exemple en remplaçant l'expression  $1 + 1$  par l'expression  $2$ , ou en factorisant certains termes, etc. On définit ensuite un ensemble de règles de simplification. Comme précédemment, on peut diviser ces règles en différentes catégories : suppression d'espèces intermédiaires, des réactions qui sont combinaisons d'autres réactions, fusion d'espèces symétriques, etc. On étudie alors les propriétés de nos règles, et on illustre leur application pour simplifier le réseau Tet-On. On s'intéresse ensuite à des règles de simplifications de systèmes d'équations, dérivées des règles pour les réseaux. Enfin, on discute de notre simplification et on la compare aux simplifications existantes dans la littérature.

Ce chapitre reprend les résultats de l'article ([Madelaine et al., 2015](#)), ainsi que de ([Madelaine, Lhoussaine, et al., 2016](#)). Il les étend en proposant d'autres règles de simplification (notamment pour la symétrie) et en étudiant la simplification de systèmes.

On rappelle qu'un réseau cinétique  $M = R \& C$  est composé d'un ensemble de réactions cinétiques  $R$  et d'une contrainte  $C$ , et qu'on peut en dériver un système d'équations  $E(M)$ . Deux réseaux  $M$  et  $M'$  sont  $(\Omega, \mathcal{I})$ -congruents,  $M \sim_{\Omega, \mathcal{I}} M'$  si, pour tout contexte  $M''$  tel que  $\text{Espèces}(M'') \cap \mathcal{I} = \emptyset$ , leurs systèmes d'équations ont les mêmes solutions observables :  $\text{sol}_{\Omega}(E(M \mid M'')) = \text{sol}_{\Omega}(E(M' \mid M''))$ .

## 4.1 Similarité

On commence par définir une notion de similarité pour les expressions et les contraintes, ainsi que les systèmes d'équations et les réseaux, permettant de les simplifier en utilisant les règles usuelles de l'arithmétique. Par exemple, on peut remplacer  $1 + 1$  par  $2$ , on peut factoriser  $k_1x_A + k_2x_A$  en  $(k_1 + k_2)x_A$ , etc.

Cette notion permet d'une part de simplifier les exemples et d'avoir des expressions cinétiques plus simples. Elle permet par exemple de retrouver la cinétique classique lors de la simplification de Michaelis-Menten. Elle est aussi nécessaire pour appliquer certaines règles, lorsqu'il y a des conditions sur les cinétiques. Par exemple, pour pouvoir dire que l'expression  $x_A^2/x_A \equiv x_A$  est bien linéaire en  $x_A$ . Enfin, elle est indispensable pour pouvoir prouver la confluence dans le Chapitre 5.

La similarité  $\equiv$  pour les expressions et les contraintes est la plus petite congruence contenant les axiomes présentés dans la Fig. 4.1.

Les règles de similarité pour les expressions sont les axiomes habituels d'un corps : associativité, commutativité, distributivité, élément neutre, etc. Par exemple, l'axiome d'associativité de l'addition dit qu'on peut réécrire  $(e_1 + e_2) + e_3$  en  $e_1 + (e_2 + e_3)$ . Ces transformations sont valables uniquement sous certaines contraintes. Par exemple, l'axiome de l'inverse pour  $*$  dit qu'on peut réécrire  $e/e$  en  $1$ , à condition que la contrainte vérifie  $e \neq 0$ .

La similarité préserve l'interprétation.

**Lemme 6.** Si  $C \vdash e_1 \equiv e_2$ , alors pour toute interprétation  $\alpha$  telle que  $\llbracket C \rrbracket_\alpha = \text{true}$ , on a  $\llbracket e_1 \rrbracket_\alpha = \llbracket e_2 \rrbracket_\alpha$ .

On étend ensuite la notion aux contraintes (Fig. 4.1). Les quatre premiers axiomes propagent la similarité des expressions. Les deux axiomes suivants sont les règles classiques de la logique. Les axiomes suivants présentent les contraintes vérifiées par les nombres réels, les paramètres et les variables. On donne ensuite un ensemble de règles de similarité pour les contraintes d'égalité, de valeur constante, supérieure ou différente de zéro. Par exemple, la somme de deux expressions constantes est toujours une expression constante.

Deux contraintes similaires sont vérifiées par les mêmes interprétations.

**Lemme 7.** Si  $C_1 \equiv C_2$ , alors pour toute interprétation  $\alpha$ , on a  $\llbracket C_1 \rrbracket_\alpha = \llbracket C_2 \rrbracket_\alpha$ .

**Exemple 31.** Par exemple, considérons l'expression  $e$  et la contrainte  $C$  suivantes :

$$e = k_1k_3x_Sx_E^0/(k_2 + k_3) \quad C = (x_C^0 = k_1x_Sx_E^0/(k_2 + k_3)).$$

On verra plus tard qu'il s'agit de la contrainte obtenue lors de la simplification du réseau cinétique de Michaelis-Menten (Fig. 2.19), et de l'expression cinétique de l'unique réaction  $S \rightarrow E$ ;  $e$  restante dans ce réseau.

## Expressions

(Associativité de +)	$C \vdash (e_1 + e_2) + e_3 \equiv e_1 + (e_2 + e_3)$
(Associativité de *)	$C \vdash (e_1 * e_2) * e_3 \equiv e_1 * (e_2 * e_3)$
(Commutativité de +)	$C \vdash e_1 + e_2 \equiv e_2 + e_1$
(Commutativité de *)	$C \vdash e_1 * e_2 \equiv e_2 * e_1$
(Distributivité)	$C \vdash e * (e_1 + e_2) \equiv e * e_1 + e * e_2$
(Neutre pour +)	$C \vdash e + 0 \equiv e$
(Neutre pour *)	$C \vdash e * 1 \equiv e$
(Inverse pour +)	$C \vdash e + (-e) \equiv 0$ si $e$ est sans division
(Inverse pour *)	$C \vdash e/e \equiv 1$ si $C \equiv (C' \wedge e \neq 0)$
(Loi pour -)	$C \vdash -(-e) \equiv e$
(Loi pour /)	$C \vdash 1/(1/e) \equiv e$ si $C \equiv (C' \wedge e \neq 0)$
(Nombres réels)	$C \vdash e \equiv a$ si $a \in \mathbb{R}$ et $\forall \alpha, \beta. \llbracket e \rrbracket_\alpha = a$
(Égalité)	$C \vdash e \equiv e'$ si $C \equiv (C' \wedge e = e')$

## Contraintes

$\frac{C \vdash e \equiv e'}{C \wedge e = e'' \equiv C \wedge e' = e''}$	$\frac{C \vdash e \equiv e'}{C \wedge \text{cst}(e) \equiv C \wedge \text{cst}(e')}$			
$\frac{C \vdash e \equiv e'}{C \wedge e \neq 0 \equiv C \wedge e' \neq 0}$	$\frac{C \vdash e \equiv e'}{C \wedge e \geq 0 \equiv C \wedge e' \geq 0}$			
$\frac{}{C \wedge C \equiv C}$	$\frac{}{C \wedge \top \equiv C}$			
$\frac{a \in \mathbb{R}}{\text{cst}(a) \equiv \top}$	$\frac{a \in \mathbb{R}_{>0}}{a > 0 \equiv \top}$	$\frac{k \in \text{Param}}{\text{cst}(k) \equiv \top}$	$\frac{k \in \text{Param}}{k \neq 0 \equiv \top}$	$\frac{x \in \text{Vars}}{x \geq 0 \equiv \top}$
$\frac{}{e = e' \equiv e - e' = 0}$	$\frac{e \text{ est une expression sans division}}{e = e \equiv \top}$			
$\frac{}{e_1 = e'_1 \wedge e_2 = e'_2 \equiv e_1 = e'_1 \wedge e_1 + e_2 = e'_1 + e'_2}$	$\frac{}{e_1 = e'_1 \wedge e_2 = e'_2 \wedge e_1 \neq 0 \equiv e_1 = e'_1 \wedge e_1 e_2 = e'_1 e'_2 \wedge e_1 \neq 0}$			
$\frac{}{e_1 = e'_1 \wedge e_1 \neq 0 \equiv 1/e_1 = 1/e'_1}$	$\frac{}{e = e' \equiv -e = -e'}$			
$\frac{}{\text{cst}(e) \wedge \text{cst}(e') \equiv \text{cst}(e) \wedge \text{cst}(e + e')}$	$\frac{}{\text{cst}(e) \equiv \text{cst}(-e)}$			
$\frac{}{\text{cst}(e) \wedge \text{cst}(e') \equiv \text{cst}(e) \wedge \text{cst}(e') \wedge \text{cst}(ee')}$	$\frac{}{\text{cst}(e) \wedge e \neq 0 \equiv \text{cst}(1/e)}$			
$\frac{}{e \geq 0 \wedge e' > 0 \equiv e \geq 0 \wedge e' > 0 \wedge e + e' > 0}$	$\frac{}{e > 0 \equiv 1/e > 0}$			
$\frac{}{e \neq 0 \wedge e' \neq 0 \equiv e \neq 0 \wedge e' \neq 0 \wedge ee' \neq 0}$	$\frac{}{e \neq 0 \equiv -e \neq 0}$			
$\frac{}{e \neq 0 \equiv 1/e \neq 0}$	$\frac{\oplus \in \{+, *\}}{e \geq 0 \wedge e' \geq 0 \equiv e \geq 0 \wedge e' \geq 0 \wedge e \oplus e' \geq 0}$			

Fig 4.1: Axiomes de similarité pour les expressions et les contraintes.

Systèmes d'équations

$$\frac{C \vdash e \equiv e'}{E \wedge \dot{x} = e \wedge C} \equiv \frac{C \equiv C'}{E \wedge C} \quad \frac{C \equiv C'}{E \wedge C} \equiv \frac{C \equiv C'}{E \wedge C'}$$

Réseaux cinétiques

$$\frac{C \vdash e \equiv e'}{R \cup \{s_1 \rightarrow s_2; e\} \& C} \equiv \frac{C \equiv C'}{R \cup \{s_1 \rightarrow s_2; e'\} \& C} \quad \frac{C \equiv C'}{R \& C} \equiv \frac{C \equiv C'}{R \& C'}$$

Fig 4.2: Axiomes de similarité pour les systèmes d'équations et les réseaux cinétiques.

On veut alors réécrire cette expression pour obtenir la forme classique de la cinétique de Michaelis-Menten. On procède ici étape par étape pour illustrer la similarité. On commence par réécrire la contrainte :

$$\begin{aligned} C &= (x_C^0 = k_1 x_S x_E^0 / (k_2 + k_3)) \\ &\equiv (x_C^0 = k_1 x_S x_E^0 / (k_2 + k_3)) \wedge \top \\ &\equiv (x_C^0 = k_1 x_S x_E^0 / (k_2 + k_3)) \wedge (x_E^0 = x_E^0) \\ &\equiv (x_C^0 + x_E^0 = k_1 x_S x_E^0 / (k_2 + k_3) + x_E^0) \wedge (x_E^0 = x_E^0) \\ &\equiv (x_C^0 + x_E^0 = (k_1 x_S / (k_2 + k_3) + 1) x_E^0) \\ &\equiv ((x_C^0 + x_E^0)(k_2 + k_3) = (k_1 x_S + k_2 + k_3) x_E^0) \\ &\equiv ((x_C^0 + x_E^0)(k_2 + k_3) = (k_1 x_S + k_2 + k_3) x_E^0) \wedge (k_1 x_S + k_2 + k_3 \neq 0) \\ &\equiv ((x_C^0 + x_E^0)(k_2 + k_3) / (k_1 x_S + k_2 + k_3) = x_E^0) \end{aligned}$$

On peut alors utiliser la contrainte pour réécrire l'expression, en remplaçant  $x_E^0$  par la valeur calculée ci-dessus :

$$\begin{aligned} C \vdash e &\equiv \frac{k_1 k_3 x_S (x_C^0 + x_E^0) (k_2 + k_3)}{(k_1 x_S + k_2 + k_3) (k_2 + k_3)} \\ C \vdash e &\equiv \frac{k_1 k_3 x_S (x_C^0 + x_E^0)}{(k_1 x_S + k_2 + k_3)} \\ C \vdash e &\equiv k_3 (x_C^0 + x_E^0) \frac{x_S}{((k_2 + k_3) / k_1 + x_S)}. \end{aligned}$$

On retrouve bien la forme classique de la cinétique de Michaelis-Menten :

$$e \equiv V \frac{x_S}{K + x_S}, \text{ avec } V = k_3 (x_C^0 + x_E^0) \text{ et } K = \frac{k_2 + k_3}{k_1}.$$

On étend la notion de similarité de façon naturelle aux systèmes d'équations et aux réseaux cinétiques, comme illustré dans la Fig. 4.2.

La similarité préserve la congruence.

**Lemme 8.**  $M \equiv M'$  implique  $M \sim M'$ .

## 4.2 Règles de simplification

On présente ici un ensemble de règles de simplification préservant la sémantique déterministe des réseaux de réactions cinétiques. On peut partitionner ces règles dans les mêmes catégories que dans le chapitre précédent. La première catégorie de règles simplifie une espèce intermédiaire, c'est-à-dire une espèce interne et à l'équilibre. La seconde catégorie ne contient qu'une seule règle, qui simplifie une réaction si elle est la combinaison linéaire d'autres réactions du réseau. Notons dans ce cas que les expressions cinétiques de ces autres réactions devront alors être modifiées. On présente ensuite des règles de simplification basées sur les symétries entre espèces. Enfin, la dernière catégorie regroupe les autres règles.

### 4.2.1 Simplification d'une espèce intermédiaire

On cherche ici à supprimer des espèces intermédiaires. Une espèce  $A$  est *intermédiaire* si elle est interne ( $A \in \mathcal{I}$ ) et à l'équilibre ( $cst(x_A)$ ). On peut alors, sous certaines conditions, supprimer complètement  $A$  du réseau. Pour cela, on va combiner deux-à-deux les réactions produisant du  $A$  avec les réactions consommant du  $A$ . La difficulté est alors de calculer les cinétiques des nouvelles réactions. En particulier, comme  $A$  peut aussi être un modificateur dans une réaction, il est nécessaire de calculer sa concentration  $x_A$ . On présente tout d'abord les règles de façon informelles, avant de détailler leur fonctionnement plus précisément.

Dans la première règle, (**INTERMÉDIAIRE - LINÉAIRE**), on suppose que les expressions cinétiques  $e$  des réactions produisant du  $A$  ne dépendent pas de  $x_A$  :  $x_A \notin \text{Vars}(e)$ . Les expressions des réactions consommant du  $A$  sont linéaires en  $x_A$ , c'est-à-dire de forme  $x_A e'$ , avec  $x_A \notin \text{Vars}(e')$ . Dans ce cas, l'équation différentielle ordinaire pour  $A$  sera  $\dot{x}_A = T - x_A T'$ , avec  $T$  et  $T'$  des expressions telles que  $x_A \notin \text{Vars}(T), \text{Vars}(T')$ . Le réseau initial doit de plus vérifier les contraintes  $cst(x_A)$  et  $T' \neq 0$  (pour tout temps  $t$ ). On peut donc facilement calculer la valeur  $x_A = T/T'$ .

Dans la règle (**INTERMÉDIAIRE - INVERSIBLE**), il existe cette fois une expression  $e_A$ , avec  $x_A \in \text{Vars}(e_A)$ , telle que les expressions cinétiques des réactions consommant du  $A$  soient toutes de forme  $e_A e$ , avec  $x_A \notin \text{Vars}(e)$ . Ainsi, on obtient à partir de l'ODE de  $A$  l'équation  $\dot{x}_A = T/T'$ . En imposant que la fonction  $e_A$  soit *inversible* pour  $x_A$ , on peut alors calculer la valeur de  $x_A$ . Notons que (**INTERMÉDIAIRE - LINÉAIRE**) est un cas particulier de cette règle, avec  $e_A = x_A$  trivialement inversible.

Enfin, dans la règle (**INTERMÉDIAIRE - DÉPENDANTE**), on suppose encore qu'il existe une expression  $e_A e$  avec  $x_A \notin \text{Vars}(e)$ , et que toutes les expressions impliquées dans le réseau soient *dépendantes* de  $A$  en  $e_A$ , ce qui implique que la valeur de  $x_A$  n'est pas nécessaire pour simplifier le réseau, la valeur de  $e_A$  est suffisante.

(INTERMÉDIAIRE - LINÉAIRE)

$$\begin{array}{c}
A \in \mathcal{I} \quad \text{split}(R, A, x_A) = (R_{prod}, R_{reac}, R_{mod}) \\
T = \sum_{s_1 \rightarrow s_2 + aA; e \in R_{prod}} ae \quad T' = \sum_{s'_1 + a'A \rightarrow s'_2; e' \in R_{reac}} d'e' \quad F = T/T' \\
C \equiv C' \wedge cst(x_A) \wedge T' \neq 0 \\
\hline
R\&C \Rightarrow (R_{prod} \diamond_{T'} R_{reac}) \cup R_{mod}[x_A := F] \\
&\&C[x_A := F] \wedge x_A^0 = F
\end{array}$$

(INTERMÉDIAIRE - INVERSIBLE)

$$\begin{array}{c}
A \in \mathcal{I} \quad \text{split}(R, A, e_A) = (R_{prod}, R_{reac}, R_{mod}) \\
T = \sum_{s_1 \rightarrow s_2 + aA; e \in R_{prod}} ae \quad T' = \sum_{s'_1 + a'A \rightarrow s'_2; e' \in R_{reac}} d'e' \quad F = T/T' \\
C \equiv C' \wedge cst(x_A) \wedge T' \neq 0 \\
e_A \text{ est inversible pour } x_A \text{ et son inverse est } e_A^{-1} \\
\hline
R\&C \Rightarrow (R_{prod} \diamond_{T'} R_{reac}) \cup R_{mod}[x_A := e_A^{-1}[x_A := F]] \\
&\&C[x_A := e_A^{-1}[x_A := F]] \wedge x_A^0 = e_A^{-1}[x_A := F]
\end{array}$$

(INTERMÉDIAIRE - DÉPENDANTE)

$$\begin{array}{c}
A \in \mathcal{I} \quad \text{split}(R, A, e_A) = (R_{prod}, R_{reac}, R_{mod}) \\
T = \sum_{s_1 \rightarrow s_2 + aA; e \in R_{prod}} ae \quad T' = \sum_{s'_1 + a'A \rightarrow s'_2; e' \in R_{reac}} d'e' \quad F = T/T' \\
C \equiv C' \wedge cst(x_A) \wedge T' \neq 0 \\
\forall e \in \text{Expr}(R\&C). \text{dep}(e, x_A, e_A) \\
\hline
R\&C \Rightarrow (R_{prod} \diamond_{T'} R_{reac}) \cup R_{mod}[e_A := F] \\
&\&C'[e_A := F] \wedge T' \neq 0 \wedge e_A[x_A := x_A^0] = F
\end{array}$$

Fig 4.3: Règles de simplification d'espèce intermédiaire en déterministe, avec  $\mathcal{I}$  l'ensemble des espèces internes.

On définit les notions d'inversibilité et de dépendance suivante.

**Définition 15** (Inverse). Soit  $x$  une variable et  $e$  une expression. On dit que  $e$  est *inversible* pour  $x$  s'il existe une expression, appelée *inverse* de  $e$  et notée  $e^{-1}$ , telle que :

$$e[x := e^{-1}] \equiv x \equiv e^{-1}[x := e].$$

**Exemple 32.** Par exemple,  $3x+2$  est inversible pour  $x$ , et son inverse est  $(x-2)/3$ .

L'expression  $V \frac{x}{K+x}$  est inversible pour  $x$ , et son inverse est  $\frac{Kx}{V-x}$ . Trivialement,  $x$  est inversible pour  $x$ , et son inverse est  $x$ .

**Définition 16** (Dépendance). Soit  $A$  une espèce moléculaire et  $e, e'$  deux expressions. On dit que  $e'$  dépend de  $x_A$  en  $e$ , noté  $dep(e', x_A, e)$ , si les variables  $x_A$  et  $x_A^0$  n'apparaissent dans  $e'$  que au sein de l'expression  $e$ . Formellement, il existe une expression  $e''$  et deux variables  $y, z$  telle que :

- $e' \equiv e''[y := e][z := e^0]$ ,
- $x_A \in Vars(e), e^0 = e[x_A := x_A^0]$
- $x_A, x_A^0 \notin Vars(e'')$ ,
- $y, z \notin Vars(e), Vars(e')$ .

Dans ce cas, on définit la substitution de  $e$  par une autre expression  $e'''$  dans  $e'$  par  $e'[e := e'''] = e''[y := e''']$ .

**Exemple 33.** Par exemple, les expressions suivantes dépendent de  $x_A$  en  $x_A^2$  :  $e_1 = x_A^2 + x_B$ ,  $e_2 = x_A^2 / (k + x_A^2)$  et  $e_3 = (x_A^0)^2 + x_A^2$ . Pour  $e_3$  par exemple, on a  $e'' = y + z$ ,  $e = x_A$  et  $e^0 = x_A^0$ . Par contre, l'expression  $x_A^2 + x_A$  ne dépend pas de  $x_A$  en  $x_A^2$ .

**Définition 17.** Dans chacune de ces règles, décrites dans la Fig. 4.3, on partitionne (si c'est possible) les réactions du réseau, en fonction de l'expression  $e_A$  (pour le cas linéaire,  $e_A = x_A$ ). On définit la fonction  $split(R, A, e_A) = (R_{reac}, R_{prod}, R_{mod})$  telle que :

- $R = R_{prod} \cup R_{reac} \cup R_{mod}$ ,
- l'ensemble  $R_{prod}$  contient les réactions produisant du  $A$ . Elles doivent être de forme  $r = s_1 \rightarrow s_2 + aA$  ;  $e$ , avec  $A \notin s_1, s_2, a > 0$  et  $x_A \notin Vars(e)$ ,
- l'ensemble  $R_{reac}$  contient les réactions consommant du  $A$ . Elles doivent être de forme  $r' = s'_1 + a'A \rightarrow s'_2$  ;  $e_A e'$ , avec  $A \notin s'_1, s'_2, a' > 0$  et  $x_A \notin Vars(e')$ ,
- l'ensemble  $R_{mod}$  contient les réactions avec  $A$  comme modificateur. Elles doivent être de forme  $r = s_1 \rightarrow s_2$  ;  $e$ , avec  $A \notin s_1, s_2$  et  $x_A \in Vars(e)$ .

On définit alors les expressions suivantes, qu'on va utiliser pour calculer les nouvelles expressions cinétiques. L'expression  $T$  est la somme des cinétiques des réactions produisant du  $A$ , pondérées par la stœchiométrie :

$$T = \sum_{s_1 \rightarrow s_2 + aA; e \in R_{prod}} ae.$$



L'expression  $T'$  est la somme des cinétiques des réactions consommant du  $A$ , modulo l'expression  $e_A$  et pondérées par la stœchiométrie :

$$T' = \sum_{s'_1 + a'A \rightarrow s'_2; e_A e' \in R_{\text{reac}}} a' e'.$$

Enfin,  $F$  est le quotient des deux expressions précédentes :

$$F = T/T'.$$

Avec les conditions de la partition ci-dessus, on a directement  $x_A \notin \text{Vars}(T)$ ,  $\text{Vars}(T')$ ,  $\text{Vars}(F)$ . L'équation différentielle pour  $A$  est alors de forme  $\dot{x}_A = T - e_A T'$ . Comme de plus on a  $\text{cst}(x_A)$  et  $T' \neq 0$ , on en déduit  $e_A = F$ .

**Définition 18.** Pour toute réaction de production  $r = s_1 \rightarrow s_2 + aA$ ;  $e$  et toute réaction de consommation  $r' = s'_1 + a'A \rightarrow s'_2$ ;  $e'e_A$ , on définit la *réaction combinée par rapport à  $T'$*  par :

$$r \diamond_{T'} r' = a' s_1 + a s'_1 \rightarrow a' s_2 + a s'_2; ee'/T'.$$

On note alors  $R_{\text{prod}} \diamond_{T'} R_{\text{reac}} = \{r \diamond_{T'} r' \mid r \in R_{\text{prod}}, r' \in R_{\text{reac}}\}$  l'ensemble des combinaisons des réactions de production et de consommation.

Notons que la condition  $T' \neq 0$  implique  $R_{\text{reac}} \neq \emptyset$ . En revanche, on peut avoir  $R_{\text{prod}} = \emptyset$ . Dans ce cas,  $T = 0$  et  $x_A = 0$ . Les règles de simplification ci-dessus suppriment alors les réactions de  $R_{\text{cons}}$ , qui ne sont jamais applicables puisque  $x_A = 0$ .

**Cas linéaire** Dans la première règle, (**INTERMÉDIAIRE - LINÉAIRE**), on a  $e_A = x_A$ . L'*ODE* permet de calculer  $x_A = F$ . On combine les réactions de production et de consommation, et on remplace  $x_A$  par  $F$  dans les autres réactions et dans la contrainte. On ajoute la contrainte  $x_A^0 = F$ .

**Exemple 34.** On considère le réseau  $M_{\text{MM}}$  de l'Ex. 19 :

$$M_{\text{MM}} = \{r_1, r_2, r_3\} \& \text{cst}(x_E) \wedge \text{cst}(x_C),$$

avec  $r_1 = S + E \rightarrow C$ ;  $k_1 x_S x_E$ ,  $r_2 = C \rightarrow S + E$ ;  $k_2 x_C$  et  $r_3 = C \rightarrow P + E$ ;  $k_3 x_C$ . On veut supprimer l'espèce interne  $C$ . On a  $\text{split}(R, C, x_C) = (\{r_1\}, \{r_2, r_3\}, \emptyset)$ . On peut donc appliquer la simplification (**INTERMÉDIAIRE - LINÉAIRE**). On a  $T = k_1 x_S x_E$ ,  $T' = k_2 + k_3$  et  $F = \frac{k_1 x_S x_E}{k_2 + k_3}$ . Comme  $T'$  est une somme de paramètres, on peut directement par similarité ajouter la contrainte  $T' \neq 0$ . Après simplification, on obtient le réseau décrit à gauche de la Fig. 4.4. La contrainte  $\text{cst}(x_C)$  du réseau initial est remplacée par  $\text{cst}(F)$ , qui peut être simplifiée par similarité en  $\text{cst}(x_S)$ .

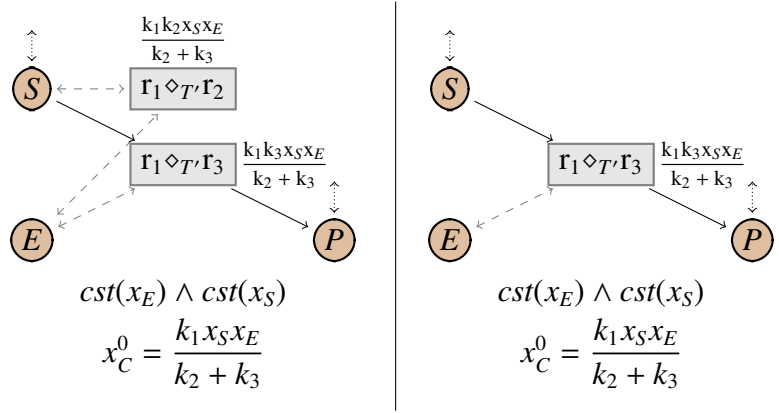
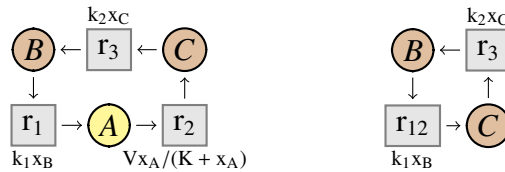
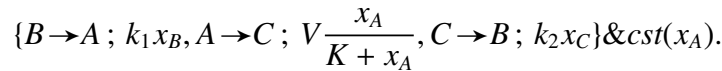


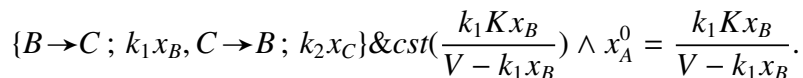
Fig 4.4: Graphes des réseaux cinétiques intermédiaire lors de la simplification de Michaelis-Menten.

**Cas inversible** Pour la règle (**INTERMÉDIAIRE - INVERSIBLE**), on impose que  $e_A$  soit inversible. On a alors  $e_A = F$  et  $x_A = e_A^{-1}[x_A := F]$ . Donc, comme précédemment, on combine les réactions de production et de consommation, on remplace  $x_A$  par  $e_A^{-1}[x_A := F]$ , et on ajoute la contrainte  $x_A^0 = e_A^{-1}[x_A := F]$ .

**Exemple 35** (Cas Inversible). On considère le réseau suivant, avec les réactions illustrées ci-dessous (à gauche) :



On veut supprimer l'espèce intermédiaire  $A$ . On a  $e_A = \frac{V x_A}{K + x_A}$ . Cette expression est inversible, avec  $e_A^{-1} = \frac{K x_A}{V - x_A}$ . On peut donc obtenir le réseau simplifié suivant, dont les réactions sont illustrées ci-dessus à droite :



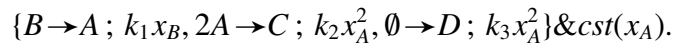
Toute interprétation  $\alpha$  du réseau simplifié doit vérifier  $V - k_1 x_B \neq 0$  pour être bien définie. C'est également le cas dans le réseau initial. En effet,  $\alpha$  vérifie

l'ODE de  $A$  à l'équilibre, donc  $0 = k_1 x_B - V \frac{x_A}{K + x_A}$ . Comme les conditions de la règle impose aussi  $V \frac{x_A}{K + x_A} > 0$ , on a  $k_1 x_B \neq 0$ . L'ODE peut se réécrire en  $k_1 x_B K = (V - k_1 x_B) x_A$ , d'où  $V - k_1 x_B \neq 0$  (on a  $K \neq 0$  par définition des paramètres).

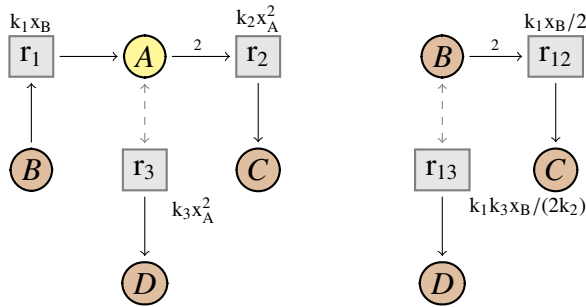
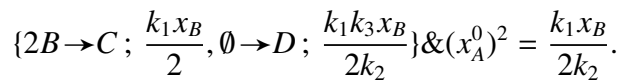
Cet exemple illustre également le fait que lorsqu'il n'y a qu'une seule réaction de consommation de  $A$  (ici  $r_2$ ), les réactions combinées ont directement la même cinétique que les réactions de production. En effet, on a  $ee'/T' = ee'/e' \equiv e$  (donc ici  $kin(r_{12}) = kin(r_1)$ ).

**Cas dépendant** Enfin, dans (**INTERMÉDIAIRE - DÉPENDANTE**), toutes les expressions des réactions avec  $A$  comme modificateur doivent être dépendantes de  $A$  en  $e_A$ . De même, la contrainte  $C$  doit être de forme  $C' \wedge cst(x_A)$ , avec  $C'$  dépendante de  $A$  en  $e_A$ . On n'a donc pas besoin de connaître la valeur de  $x_A$ . À partir de l'ODE de  $A$ , on peut calculer la valeur de  $e_A = F$ . On combine donc les réactions de production et de dégradation, et on remplace  $e_A$  par  $F$  dans les autres réactions et dans  $C'$ . Enfin, dans le réseau initial, toute interprétation  $\alpha$  doit vérifier  $cst(x_A)$  et l'ODE de  $A$ . Pour avoir les mêmes solutions, on ajoute la contrainte  $e_A[x_A := x_A^0] = F$  dans le réseau simplifié.

**Exemple 36** (Cas dépendant). On considère le réseau suivant, représenté à gauche ci-dessous :



Les expressions sont toutes dépendantes en  $x_A^2$ , qui n'est pas inversible. On peut appliquer (**INTERMÉDIAIRE - DÉPENDANTE**) et obtenir :



$$\begin{array}{c}
 \text{(COMBINAISON)} \\
 \frac{\exists a_1, \dots, a_k \in \mathbb{N}. \text{stoic}_r = \sum_{1 \leq i \leq k} a_i \text{stoic}_{r_i}}{\{r\} \cup \{r_1, \dots, r_k\} \& \top \Rightarrow \begin{array}{l} \{s_i \rightarrow s'_i; e_i + a_i \text{kin}(r) \\ | r_i = s_i \rightarrow s'_i; e_i\} \& \top \end{array}}
 \end{array}$$

Fig 4.5: Règle de simplification de combinaison en déterministe, avec  $\Omega$  la fonction d'observation et  $\mathcal{I}$  les espèces internes.

## 4.2.2 Combinaison

On considère maintenant une règle supprimant une réaction lorsqu'elle est la combinaison linéaire d'autres réactions, comme pour la règle (**DÉCOMPOSITION - SUPPRESSION**) pour le cas non déterministe. Il faut néanmoins considérer ici explicitement ces autres réactions, et modifier leurs expressions cinétiques.

La règle (**COMBINAISON**) est présentée dans la Fig. 4.5. Une réaction  $r$  est la combinaison d'un ensemble de réactions  $\{r_1, \dots, r_k\}$  s'il existe des entiers  $a_i > 0$  tels que  $\text{stoic}_r = \sum_{1 \leq i \leq k} a_i \text{stoic}_{r_i}$ . Cette règle permet notamment de supprimer les réactions cycliques de forme  $s \rightarrow s; e$ , lorsque l'ensemble de réactions est vide :  $\{r_1, \dots, r_k\} = \emptyset$ .

**Exemple 37.** On reprend le réseau simplifié de la Fig. 4.4 (gauche), après application de la règle (**INTERMÉDIAIRE - LINÉAIRE**). On peut utiliser la règle (**COMBINAISON**) pour supprimer la réaction cyclique  $\emptyset \rightarrow \emptyset; \frac{k_1 k_2 x_S x_E}{k_2 + k_3}$ . On obtient le réseau de droite.

## 4.2.3 Symétrie

On présente ici deux règles permettant de simplifier des espèces ayant un comportement symétrique. Les règles sont présentées dans la Fig. 4.6.

**Symétrie totale** Dans la première règle, (**SYMÉTRIE**), deux espèces  $A_1$  et  $A_2$  vont toujours avoir la même concentration au cours du temps, et on va donc pouvoir en supprimer une. Formellement, on peut décomposer l'ensemble des réactions en trois ensembles disjoints :  $R_{A_1}$ , celles qui utilisent  $A_1$  mais pas  $A_2$ ,  $R_{A_2}$ , les réactions qui utilisent  $A_2$  mais pas  $A_1$ , et  $R$ , les réactions pouvant avoir  $A_1$  et  $A_2$  comme modificateur. De plus, ces deux premiers ensembles sont symétriques :  $R_{A_2} = R_{A_1}[x_{A_1} := x_{A_2}][A_1 := A_2]$ . Comme  $A_1$  et  $A_2$  sont internes, et ont la même concentration initiale, on va toujours avoir  $x_{A_1} = x_{A_2}$ . On peut donc remplacer  $x_{A_2}$

(SYMÉTRIE)

$$\begin{array}{c}
A_1, A_2 \in \mathcal{I} \quad \Omega(A_2) = \mathbb{O} \quad C \equiv C' \wedge x_{A_1}^0 = x_{A_2}^0 \\
A_2 \notin \text{Espèces}(R_{A_1}) \quad A_1 \notin \text{Espèces}(R_{A_2}) \\
R_{A_2} = R_{A_1}[x_{A_1} := x_{A_2}][A_1 := A_2] \\
\forall r \in R. \text{stoic}_r(A_1) = \text{stoic}_r(A_2) = 0 \\
\hline
R_{A_1} \cup R_{A_2} \cup R\&C \Rightarrow R_{A_1} \cup ((R_{A_2}[x_{A_2} := x_{A_1}]) \setminus A_2) \cup R[x_{A_2} := x_{A_1}] \\
\&C[x_{A_2} := x_{A_1}]
\end{array}$$

(SYMÉTRIE - RÉACTANT)

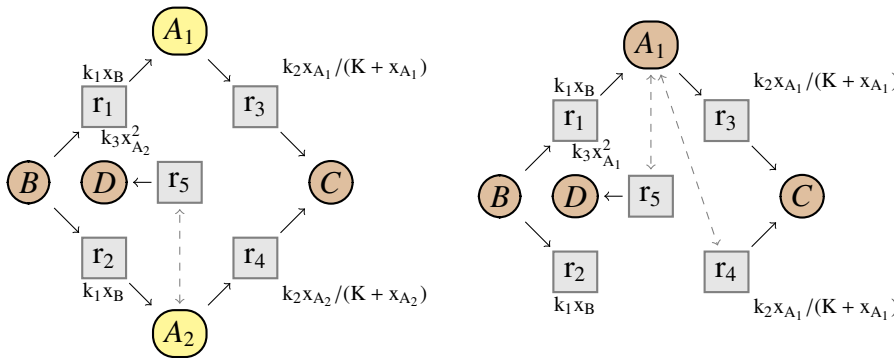
$$\begin{array}{c}
A_1, A_2, A \in \mathcal{I} \quad \Omega(A_1) = \Omega(A_2) = \Omega(A) = \mathbb{O} \\
A \notin \text{Espèces}(R_{A_1} \cup R_{A_2} \cup R\&C) \\
A_2 \notin \text{Espèces}(R_{A_1}) \quad A_1 \notin \text{Espèces}(R_{A_2}) \\
R_{A_2} = R_{A_1}[x_{A_1} := x_{A_2}][A_1 := A_2] \\
\forall r \in R_{A_1}. \exists e. \text{kin}(r) = x_{A_1}e \wedge x_{A_1} \notin \text{Vars}(e) \wedge \text{stoic}_r(A_1) < 0 \\
\forall r = s_1 \rightarrow s_2; e \in R. A_1, A_2 \notin s_1 \\
\forall e \in \text{Expr}(R) \cup \text{Expr}(C). \text{dep}(e, x_{A_1}, x_{A_1} + x_{A_2}) \wedge \text{dep}(e, x_{A_2}, x_{A_1} + x_{A_2}) \\
\hline
R_{A_1}[x_{A_1} := x_A][A_1 := A] \\
R_{A_1} \cup R_{A_2} \cup R\&C \Rightarrow \cup R[x_{A_1} + x_{A_2} := x_A][A_1 := A][A_2 := A] \\
\&C[x_{A_1} + x_{A_2} := x_A] \wedge x_A^0 = x_{A_1}^0 + x_{A_2}^0
\end{array}$$

Fig 4.6: Règles de simplification d'espèces symétriques en déterministe, avec  $\Omega$  la fonction d'observation et  $\mathcal{I}$  les espèces internes.

par  $x_{A_1}$  partout. Comme l'espèce  $A_2$  est non observable, on la supprime complètement du réseau.

**Exemple 38.** Par exemple, dans le réseau suivant, dont les réactions sont représentées ci-dessous (à gauche) :

$$R \& x_{A_1}^0 = x_{A_2}^0.$$

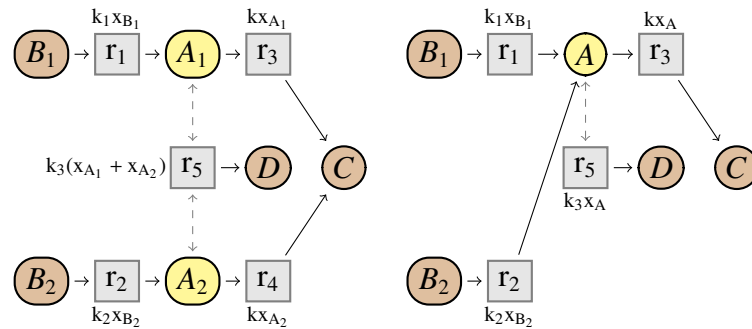


Les espèces  $A_1$  et  $A_2$  sont symétriques. En supposant que  $A_2$  n'est pas observable, on peut la supprimer, et obtenir le réseau dont les réactions sont représentées ci-dessus (à droite). Les ensembles correspondant à la règle sont  $R_{A_1} = \{r_1, r_3\}$ ,  $R_{A_2} = \{r_2, r_4\}$ , et  $R = \{r_5\}$ .

**Symétrie partielle** L'autre règle, (**SYMÉTRIE - RÉACTANT**), fusionne deux espèces  $A_1$  et  $A_2$  si elles sont toujours consommées de la même façon. Formellement, on peut partitionner les réactions en trois ensembles :  $R_{A_1}$  et  $R_{A_2}$ , les réactions qui consomment respectivement du  $A_1$  et du  $A_2$ , et  $R$ , les réactions qui ne consomment ni  $A_1$  ni  $A_2$  ( $\forall r = s_1 \rightarrow s_2 ; e \in R. A_1, A_2 \notin s_1$ ). Les réactions de  $R_{A_1}$  n'utilisent pas  $A_2$ , et ont une cinétique linéaire en  $x_{A_1}$ , et symétriquement pour  $R_{A_2}$ . De plus,  $R_{A_2} = R_{A_1}[A_1 := A_2]$ . Dans les réactions de  $R$ , ainsi que dans la contrainte C, les concentrations  $x_{A_1}$  et  $x_{A_2}$  apparaissent toujours ensemble, sous la forme  $x_{A_1} + x_{A_2}$  (formellement,  $\forall e \in Expr(R) \cup Expr(C)$ , on a  $dep(e, x_{A_1}, x_{A_1} + x_{A_2})$  et  $dep(e, x_{A_2}, x_{A_1} + x_{A_2})$ ). Enfin,  $A_1$  et  $A_2$  ne sont pas observables.

On va alors prendre une nouvelle espèce,  $A$ , interne et non observable, telle que chaque molécule de  $A$  représente une molécule de  $A_1$  ou une molécule de  $A_2$ , et que  $x_A = x_{A_1} + x_{A_2}$ . On remplace  $A_1$  et  $A_2$  par  $A$  dans  $R_{A_1}$  et  $R$ ,  $x_{A_1}$  par  $x_A$  dans  $R_{A_1}$ ,  $x_{A_1} + x_{A_2}$  par  $x_A$  dans  $R$  et C, et on supprime  $R_{A_2}$ . Enfin, on ajoute la contrainte  $x_A^0 = x_{A_1}^0 + x_{A_2}^0$ .

**Exemple 39.** Par exemple, considérons le réseau  $R \& \top$ , donc les réactions sont représentées ci-dessous (à gauche).



En supposant que  $A_2$  n'est pas observable, on peut la supprimer. Les ensembles de réactions correspondant à ceux de la règle sont  $R_{A_1} = \{r_3\}$ ,  $R_{A_2} = \{r_4\}$ , et  $R = \{r_1, r_2, r_5\}$ . On obtient le réseau suivant, dont les réactions sont représentées ci-dessus (à droite) :

$$R' \& x_A^0 = x_{A_1}^0 + x_{A_2}^0.$$

#### 4.2.4 Autres règles

Enfin, on présente dans la Fig. 4.7 d'autres règles de simplification, n'appartenant pas aux catégories précédentes.

**Factorisation** La règle (**FACTORISATION**) factorise les réactants et les produits d'une réaction. Ainsi, une réaction de forme  $as_1 \rightarrow as_2$  ;  $e$ , avec  $a > 0$ , est remplacée par la réaction  $s_1 \rightarrow s_2$  ;  $ae$ .

**Modificateur Constant** La deuxième règle, (**MODIFICATEUR CONSTANT**), traite le cas d'un modificateur  $A$ , interne, qui n'est jamais présent ni en réactant ni en produit, et donc dont la concentration va rester constante au cours du temps. On peut donc remplacer sa concentration par sa concentration initiale.

**Exemple 40.** On reprend le réseau partiellement simplifié de Michaelis-Menten la Fig. 4.4 (droite), après application de la règle (**COMBINAISON**). On peut utiliser la règle (**MODIFICATEUR CONSTANT**) pour supprimer l'espèce  $x_E$ . On obtient une unique réaction  $S \rightarrow P$  ;  $k_1 k_3 x_S x_E^0 / (k_2 + k_3)$ , et la contrainte  $cst(x_S) \wedge x_C^0 = k_1 x_S x_E^0 / (k_2 + k_3)$ .

Par similarité, on peut réécrire la contrainte et l'expression cinétique, et obtenir le réseau simplifié  $M'_{MM}$  représenté dans la Fig. 4.8, avec la cinétique classique de Michaelis-Menten.

Notons en particulier que pour que cette simplification soit exacte, il est donc nécessaire que  $S$  soit également à l'équilibre. Dans le cas contraire,  $C$  ne pourra pas être à l'équilibre, et donc la simplification ne sera pas correcte.

$$\begin{array}{l}
\text{(FACTORISATION)} \\
\frac{\{as_1 \rightarrow as_2; e\} \& \top \Rightarrow \{s_1 \rightarrow s_2; ae\} \& \top}{} \\
\text{(MODIFICATEUR CONSTANT)} \\
\frac{A \in \mathcal{I} \quad \forall r \in R. \text{stoic}_r(A) = 0}{R \& C \Rightarrow R[x_A := x_A^0] \& C[x_A := x_A^0]} \\
\text{(ESPÈCE INUTILE)} \\
\frac{A \in \mathcal{I} \quad \Omega(A) = 0 \\
\forall s_1 \rightarrow s_2; e \in R. A \notin s_1 \wedge x_A \notin \text{Vars}(e) \\
\forall r' \in R'. \forall B \neq A. \text{stoic}_{r'}(B) = 0}{R \cup R' \& \top \Rightarrow R \setminus A \& \top} \\
\text{(RÉACTIONS INAPPLICABLES)} \\
\frac{A \in \mathcal{I} \quad R \neq \emptyset \\
\forall r \in R. \text{kin}(r) \equiv x_A e \quad \forall r' \in R'. \text{stoic}_{r'}(A) = 0}{R \cup R' \& C \wedge x_A^0 = 0 \Rightarrow R'[x_A := 0] \& C[x_A := 0] \wedge x_A^0 = 0} \\
\text{(CONTEXTE)} \\
\frac{M \Rightarrow M' \quad \text{Espèces}(M'') \cap \mathcal{I} = \emptyset}{M \mid M'' \Rightarrow M' \mid M''} \\
\text{(SIMILARITÉ)} \\
\frac{M_1 \equiv M'_1 \quad M'_1 \Rightarrow M'_2 \quad M'_2 \equiv M_2}{M_1 \Rightarrow M_2}
\end{array}$$

Fig 4.7: Autres règles de simplification en déterministe, avec  $\Omega$  la fonction d'observation et  $\mathcal{I}$  l'ensemble des espèces internes.



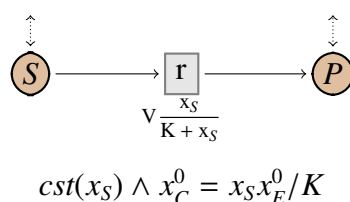


Fig 4.8: Graphe du réseau cinétique simplifié de Michaelis-Menten, avec  $V = k_3(x_C^0 + x_E^0)$  et  $K = (k_2 + k_3)/k_1$ .

**Espèce Inutile** Cette règle est similaire à la règle (**ESPÈCE INUTILE**) pour le cas non déterministe : une espèce  $A$ , interne et non observable, qui n'est jamais en réactant des réactions, sauf dans des réactions  $R'$  qui ne modifient que la quantité de  $A$ . De plus, ici sa concentration  $x_A$  n'est pas non plus dans les expressions cinétiques, ni dans les contraintes. Donc  $A$  n'est jamais nécessaire et peut être supprimée.

**Réactions Inapplicables** De façon similaire, la règle (**RÉACTIONS INAPPLICABLES**) étend la version non déterministe. Une espèce  $A$  n'est pas présente initialement, et sa concentration ne peut pas augmenter. Alors toutes les réactions avec une expression cinétique de forme  $x_A e$ , ne seront jamais applicables, et peuvent être supprimées. On remplace également  $x_A$  par 0 dans les contraintes et les autres expressions cinétiques.

**Contexte et similarité** Enfin, les deux dernières règles étendent la simplification avec la composition de réseaux ou la similarité. La règle (**CONTEXTE**) décrit le fait que si on peut simplifier un module d'un grand réseau, alors la simplification est aussi valable pour le grand réseau. La règle (**SIMILARITÉ**) précise que la simplification est en fait définie modulo la notion de similarité entre réseaux cinétiques.

## 4.2.5 Propriétés

On prouve ici les différentes propriétés de nos règles : elles simplifient bien les réseaux, et sont correctes pour la congruence déterministe.

**Simplification** On définit la taille des réseaux cinétiques de la façon suivante, similaire aux réseaux sans cinétique.

**Définition 19** (Taille et ordre des réseaux cinétiques). Soit  $M = R\&C$  un réseau cinétique. Sa *taille*  $|M|$  est le triplet  $(n, m, p)$ , avec :

- $n = \text{Espèces}(M)$  le nombre d'espèces de  $M$ ,
- $m = |R|$  le nombre de réactions,
- $p = \sum_{s_1 \rightarrow s_2; e \in R} |s_1 + s_2|$  la taille combinée de toutes les solutions du réseau.

Un réseau  $M$  de taille  $(n, m, p)$  est alors *plus petit* qu'un réseau  $M'$  de taille  $(n', m', p')$ , noté  $M < M'$ , si  $(n, m, p) < (n', m', p')$  avec l'ordre lexicographique, c'est-à-dire :

- $n < n'$ ,
- ou  $n = n'$  et  $m < m'$ ,
- ou  $n = n'$ ,  $m = m'$  et  $p < p'$ .

**Proposition 11.** La transformation  $\Rightarrow$  est une simplification.

*Proof.* Les règles d'élimination des espèces intermédiaires de la Fig. 4.3, ainsi que les règles de symétrie de la Fig. 4.6 et que les règles (**MODIFICATEUR CONSTANT**) et (**ESPÈCE INUTILE**) diminuent le nombre d'espèces moléculaires.

Les règles (**COMBINAISON**) et (**RÉACTIONS INAPPLICABLES**) diminuent le nombre de réactions sans augmenter le nombre d'espèces.

Enfin, la règle (**FACTORISATION**) diminue la taille combinée des solutions du réseau.  $\square$

**Correction** On prouve que les règles de simplification sont correctes.

**Théorème 4.** La simplification  $\Rightarrow_{\Omega, \mathcal{I}}$  est correcte pour la congruence  $\sim_{\Omega, \mathcal{I}}$ .

*Proof.* On veut montrer que si  $M \Rightarrow_{\Omega, \mathcal{I}} M'$ , alors pour tout contexte  $M''$  tel que  $\text{Espèces}(M'') \cap \mathcal{I} = \emptyset$ , on a  $M | M'' \simeq_{\Omega} M' | M''$ , c'est-à-dire  $\text{sol}_{\Omega}(E(M | M'')) = \text{sol}_{\Omega}(E(M' | M''))$ .

Commençons par la règle (**INTERMÉDIAIRE - LINÉAIRE**), appliquée sur une espèce  $A$ . On considère tout d'abord l'équation différentielle de  $A$  dans  $E(M | M'')$ . L'espèce  $A$  est interne, on a donc :

$$\begin{aligned} \dot{x}_A &= \sum_{s_1 \rightarrow s_2 + aA; e \in R_{\text{prod}}} ae - \sum_{s'_1 + a'A \rightarrow s'_2; e' \in R_{\text{reac}}} a'e' x_A \\ &= T - x_A T'. \end{aligned}$$

Toute solution  $\alpha \in \text{sol}(E(M \mid M''))$  doit vérifier cette équation, ainsi que les contraintes  $\text{cst}(x_A)$ , qui impliquent  $\dot{x}_A = 0$ , et  $T' > 0$ . Donc en particulier,  $\alpha$  doit vérifier  $x_A = T/T' = F$ . Donc, puisque  $\alpha$  vérifie C,  $\alpha$  vérifie aussi  $C[x_A := F]$  et  $x_A^0 = F$ . L'équation différentielle pour A dans le réseau simplifié est simplement  $\dot{x}_A = 0$ , et est donc bien vérifiée.

Soit  $B \neq A$  une autre espèce. On va montrer que si  $\alpha$  vérifie l'équation différentielle ordinaire de B dans  $E(M \mid M'')$ , alors elle vérifie aussi l'ODE de B dans  $E(M' \mid M'')$ . Considérons l'ODE dans  $E(M \mid M'')$ . On note  $f(M'', x_B)$  l'évolution de  $x_B$  due au contexte  $M''$  :

$$\begin{aligned} \dot{x}_B = & \sum_{s_1 \rightarrow s_2 + aA; e \in R_{\text{prod}}} (s_2(B) - s_1(B))e + \sum_{s'_1 + a'A \rightarrow s'_2; e' x_A \in R_{\text{reac}}} (s'_2(B) - s'_1(B))x_A e \\ & + \sum_{s_1 \rightarrow s_2; e \in R_{\text{mod}}} (s_2(B) - s_1(B))e + f(M'') \end{aligned}$$

Dans  $E(M' \mid M'')$ , en utilisant  $x_A = F$ , on peut réécrire l'ODE de B de la façon suivante :

$$\begin{aligned} \dot{x}_B = & \sum_{r \diamond_{T'} r' \in R_{\text{prod}} \diamond_{T'} R_{\text{reac}}} \text{stoic}_{r \diamond_{T'} r'}(B) \text{kin}(r \diamond_{T'} r) \\ & + \sum_{s_1 \rightarrow s_2; e[x_A := F] \in R_{\text{mod}}[x_A := F]} (s_2(B) - s_1(B))e[x_A := F] + f(M'') \\ = & \sum_{s_1 \rightarrow s_2 + aA; e \in R_{\text{prod}}} \sum_{s'_1 + a'A \rightarrow s'_2; e' x_A \in R_{\text{reac}}} (a' s_2(B) + a s'_2(B) - a' s_1(B) - a s'_1(B))e e' / T' \\ & + \sum_{s_1 \rightarrow s_2; e \in R_{\text{mod}}} (s_2(B) - s_1(B))e + f(M'') \\ = & \sum_{s_1 \rightarrow s_2 + aA; e \in R_{\text{prod}}} \sum_{s'_1 + a'A \rightarrow s'_2; e' x_A \in R_{\text{reac}}} (a' s_2(B) - a' s_1(B))e e' / T' \\ & + \sum_{s_1 \rightarrow s_2 + aA; e \in R_{\text{prod}}} \sum_{s'_1 + a'A \rightarrow s'_2; e' x_A \in R_{\text{reac}}} (a s'_2(B) - a s'_1(B))e e' / T' \\ & + \sum_{s_1 \rightarrow s_2; e \in R_{\text{mod}}} (s_2(B) - s_1(B))e + f(M'') \\ = & \sum_{s_1 \rightarrow s_2 + aA; e \in R_{\text{prod}}} (s_2(B) - s_1(B))e \left( \sum_{s'_1 + a'A \rightarrow s'_2; e' x_A \in R_{\text{reac}}} a' e' \right) / T' \\ & + \sum_{s'_1 + a'A \rightarrow s'_2; e' x_A \in R_{\text{reac}}} (s'_2(B) - s'_1(B))e' \left( \sum_{s_1 \rightarrow s_2 + aA; e \in R_{\text{prod}}} a e \right) / T' \\ & + \sum_{s_1 \rightarrow s_2; e \in R_{\text{mod}}} (s_2(B) - s_1(B))e + f(M'') \\ = & \sum_{s_1 \rightarrow s_2 + aA; e \in R_{\text{prod}}} (s_2(B) - s_1(B))e T' / T' \\ & + \sum_{s'_1 + a'A \rightarrow s'_2; e' x_A \in R_{\text{reac}}} (s'_2(B) - s'_1(B))e' F \end{aligned}$$

$$\begin{aligned}
& + \sum_{s_1 \rightarrow s_2; e \in R_{mod}} (s_2(B) - s_1(B))e + f(M'') \\
= & \sum_{s_1 \rightarrow s_2 + aA; e \in R_{prod}} (s_2(B) - s_1(B))e + \sum_{s'_1 + a'A \rightarrow s'_2; e' \in R_{reac}} (s'_2(B) - s'_1(B))e' x_A \\
& + \sum_{s_1 \rightarrow s_2; e \in R_{mod}} (s_2(B) - s_1(B))e + f(M'')
\end{aligned}$$

On retrouve donc la même *ODE*, donc  $\alpha$  est aussi une solution pour l'*ODE* de  $B$  dans le réseau réduit. Comme  $A$  n'est pas observable, toute solution observable de  $E(M \mid M'')$  est donc aussi une solution observable de  $E(M' \mid M'')$ .

Réciproquement, si  $\alpha'$  est une solution pour  $E(M' \mid M'')$ .  $\alpha'$  doit vérifier  $x_A^0 = F$  et l'*ODE* de  $A$  :  $\dot{x}_A = 0$ , donc  $x_A = F$ . Pour toute espèce  $B \neq A$ , avec le même raisonnement que ci-dessus,  $\alpha'$  vérifie l'*ODE* de  $B$  dans  $E(M \mid M'')$ . De même, trivialement  $\alpha'$  vérifie  $C$ . Donc  $E(M \mid M'')$  et  $E(M' \mid M'')$  ont les mêmes solutions, et donc  $M \sim_{\Omega, I} M'$ .

La preuve de la règle (**INTERMÉDIAIRE - INVERSIBLE**) est similaire, en remplaçant  $x_A$  par l'expression inversible  $e_A$ . Pour (**INTERMÉDIAIRE - DÉPENDANTE**), seules les contraintes sont légèrement différentes. Si  $\alpha$  est une solution pour réseau initial (dans son contexte), avec le même raisonnement que ci-dessus,  $\alpha$  vérifie  $e_A = F$ , et donc  $e_A[x_A := x_A^0] = F$ , puisque  $x_A$  est constante. Réciproquement, si  $\alpha'$  est une solution du réseau simplifié, elle vérifie  $e_A[x_A := x_A^0] = F$ , et donc  $e_A = F$ . Donc  $\alpha'$  est bien une solution pour le réseau initial.

Considérons la règle (**COMBINAISON**), et  $A$  une espèce. L'équation différentielle pour  $A$  dans le réseau initial dans son contexte :

$$\begin{aligned}
\dot{x}_A &= stoic_r(A)kin(r) + \sum_{1 \leq i \leq k} stoic_{r_i}(A)kin(r_i) + f(M'', x_A) \\
&= \left( \sum_{1 \leq i \leq k} stoic_{r_i}(A)kin(r) + \sum_{1 \leq i \leq k} stoic_{r_i}(A)kin(r_i) + f(M'', x_A) \right) \\
&= \sum_{1 \leq i \leq k} (stoic_{r_i}(A)kin(r) + kin(r_i)) + f(M'', x_A).
\end{aligned}$$

On retrouve donc l'*ODE* de  $A$  dans le réseau simplifié, donc les deux réseaux ont les mêmes solutions, et  $M \sim_{\Omega, I} M'$ .

Considérons la règle (**SYMÉTRIE**), et soit  $\alpha$  une solution du réseau initial dans son contexte. L'espèce  $A_1$  est interne, donc l'*ODE* pour  $A_1$  est :

$$\dot{x}_{A_1} = \sum_{r \in R_{A_1}} stoic_r(A_1)kin(r).$$

L'équation pour  $A_2$  est :

$$\dot{x}_{A_2} = \sum_{r \in R_{A_2}} \text{stoic}_r(A_2) \text{kin}(r) = \sum_{r \in R_{A_1}} \text{stoic}_r(A_1) \text{kin}(r) [x_{A_1} := x_{A_2}].$$

Les deux concentrations sont donc solutions d'équations symétriques. Et comme  $x_{A_1}^0 = x_{A_2}^0$ , on obtient  $x_{A_1} = x_{A_2}$ . On définit  $\alpha''$  par :

$$\alpha''(x_X) = \begin{cases} \alpha(x_X) & \text{si } X \neq A_2 \\ \alpha(x_{A_2}^0) & \text{pour } X = A_2. \end{cases}$$

Dans le réseau simplifié, l'*ODE* pour  $A_1$  est la même, donc est toujours vérifiée par  $\alpha''$ . L'espèce n'est pas dans le réseau simplifié, donc  $\dot{x}_{A_2} = 0$  et  $x_{A_2} = x_{A_2}^0$ . Pour toute autre espèce  $B$ , son *ODE* dans le réseau simplifié est :

$$\begin{aligned} \dot{x}_B &= \sum_{r \in R_{A_1}} \text{stoic}_r(B) \text{kin}(r) + \sum_{r \in R_{A_2}} \text{stoic}_r(B) \text{kin}(r) [x_{A_2} := x_{A_1}] \\ &+ \sum_{r \in R} \text{stoic}_r(B) \text{kin}(r) [x_{A_2} := x_{A_1}] + f(M'', x_B) \\ &= \sum_{r \in R_{A_1}} \text{stoic}_r(B) \text{kin}(r) + \sum_{r \in R_{A_2}} \text{stoic}_r(B) \text{kin}(r) \\ &+ \sum_{r \in R} \text{stoic}_r(B) \text{kin}(r) + f(M'', x_B) \end{aligned}$$

Donc  $\alpha''$  vérifie les *ODEs* du réseau simplifié, ainsi que trivialement la contrainte  $C[x_{A_2} := x_{A_1}]$ . Comme  $\Omega(A_2) = \emptyset$ , on a  $\Omega(\alpha) = \Omega(\alpha'')$  et  $\text{sol}_\Omega(M) \subseteq \text{sol}_\Omega(M')$ .

Réciproquement, pour tout solution  $\alpha'$  du réseau simplifié, on considère la solution  $\alpha''$  définie par :

$$\alpha''(x_X) = \begin{cases} \alpha'(x_{A_1}) & \text{si } X = A_2 \\ \alpha'(x_X) & \text{sinon.} \end{cases}$$

Alors  $\alpha''$  a la même observation que  $\alpha'$ , et est solution du réseau initial. Donc les deux réseaux ont les mêmes solutions observables, et  $M \sim_{\Omega, I} M'$ .

Considérons maintenant la règle (**SYMÉTRIE - RÉACTANT**), et  $\alpha$  une solution du réseau initial. On a dans le réseau initial :

$$\begin{aligned} \dot{x}_{A_1} &= \sum_{r \in R} \text{stoic}_r(A_1) \text{kin}(r) - \sum_{s_1 + a_1 A_1 \rightarrow s_2; x_{A_1} e \in R_{A_1}} a_1 x_{A_1} e \\ \dot{x}_{A_2} &= \sum_{r \in R} \text{stoic}_r(A_2) \text{kin}(r) - \sum_{s_1 + a_2 A_2 \rightarrow s_2; x_{A_2} e \in R_{A_2}} a_2 x_{A_2} e \\ &= \sum_{r \in R} \text{stoic}_r(A_2) \text{kin}(r) - \sum_{s_1 + a_1 A_1 \rightarrow s_2; x_{A_1} e \in R_{A_1}} a_1 x_{A_2} e. \end{aligned}$$

On considère alors la solution  $\alpha''$  définie par :

$$\alpha''(x_X) = \begin{cases} \alpha(x_{A_1} + x_{A_2}) & \text{si } X = A \\ \alpha(x_X) & \text{sinon.} \end{cases}$$

Cette solution  $\alpha''$  a la même observation que  $\alpha$ . On note  $R' = R[x_{A_1} + x_{A_2} := x_A][A_1 := A][A_2 := A]$  et  $R'_{A_1} = R_{A_1}[x_{A_1} := x_A][A_1 := A]$ . Alors  $\alpha''$  vérifie :

$$\begin{aligned} \dot{x}_A &= \dot{x}_{A_1} + \dot{x}_{A_2} \\ &= \sum_{r \in R} \text{stoic}_r(A_1) \text{kin}(r) - \sum_{s_1 + a_1 A_1 \rightarrow s_2; x_{A_1} e \in R_{A_1}} a_1 x_{A_1} e \\ &\quad + \sum_{r \in R} \text{stoic}_r(A_2) \text{kin}(r) - \sum_{s_1 + a_1 A_1 \rightarrow s_2; x_{A_1} e \in R_{A_1}} a_1 x_{A_2} e \\ &= \sum_{r \in R} (\text{stoic}_r(A_1) + \text{stoic}_r(A_2)) \text{kin}(r) - \sum_{s_1 + a_1 A_1 \rightarrow s_2; x_{A_1} e \in R_{A_1}} a_1 (x_{A_1} + x_{A_2}) e \\ &= \sum_{r \in R'} \text{stoic}_r(A) \text{kin}(r) - \sum_{s_1 + a_1 A \rightarrow s_2; x_C e \in R'_{A_1}} a_1 x_A e. \end{aligned}$$

Donc  $\alpha''$  vérifie l'ODE de  $A$  dans le réseau simplifié. Pour tout autre espèce  $B$ , l'ODE initiale se réécrit en :

$$\begin{aligned} \dot{x}_B &= \sum_{r = s_1 \rightarrow s_2; x_{A_1} e \in R_{A_1}} \text{stoic}_r(B) x_{A_1} e + \sum_{r = s_1 \rightarrow s_2; x_{A_2} e \in R_{A_2}} \text{stoic}_r(B) x_{A_2} e \\ &\quad + \sum_{r \in R} \text{stoic}_r(B) \text{kin}(r) + f(M'', x_B) \\ &= \sum_{r = s_1 \rightarrow s_2; x_{A_1} e \in R_{A_1}} \text{stoic}_r(B) (x_{A_1} + x_{A_2}) e + \sum_{r \in R} \text{stoic}_r(B) \text{kin}(r) + f(M'', x_B) \\ &= \sum_{r = s_1 \rightarrow s_2; x_A e \in R'_{A_1}} \text{stoic}_r(B) x_A e + \sum_{r \in R'} \text{stoic}_r(B) \text{kin}(r) + f(M'', x_B). \end{aligned}$$

Donc  $\alpha''$  vérifie aussi l'ODE du réseau simplifié. Donc  $\alpha''$  est bien une solution du réseau simplifié. Réciproquement, pour tout  $\alpha'$  solution du réseau simplifié, avec le même raisonnement que ci-dessus, on peut construire une solution  $\alpha''$  du réseau initial avec la même observation. Donc  $M \sim_{\Omega, I} M'$ .

Le règle (**FACTORISATION**) est trivialement vraie, en réécrivant l'ODE de toute espèce  $A$  de la façon suivante :

$$\begin{aligned} \dot{x}_A &= (a s_2(A) - a s_1(A)) e + f(M'', x_A) \\ &= (s_2(A) - s_1(A)) a e + f(M'', x_A). \end{aligned}$$

Considérons la règle (**MODIFICATEUR CONSTANT**). L'ODE pour  $A$  est  $\dot{x}_A = 0$  dans les deux réseaux, donc  $x_A = x_A^0$ , et les réseaux sont bien congruents.

Considérons la règle (**ESPÈCE INUTILE**). Pour toute espèce  $B \neq A$ , son équation différentielle est :

$$\begin{aligned} \dot{x}_B &= \sum_{r \in R} \text{stoic}_r(B) \text{kin}(r) + \sum_{r' \in R'} \text{stoic}_{r'}(B) \text{kin}(r') \\ &= \sum_{r \in R} \text{stoic}_r(B) \text{kin}(r). \end{aligned}$$

Comme de plus  $x_A \notin \text{kin}(r)$  pour toute réaction  $r \in R$ , l'ODE du réseau simplifié est la même. L'équation différentielle de  $A$  est modifiée, mais comme  $A$  n'est pas observable, les deux réseaux ont bien les mêmes solutions observables.

Considérons la règle (**RÉACTIONS INAPPLICABLES**). L'équation différentielle ordinaire pour  $A$  est :

$$\begin{aligned} \dot{x}_A &= \sum_{r \in R} \text{stoic}_r(A) \text{kin}(r) + \sum_{r' \in R'} \text{stoic}_{r'}(A) \text{kin}(r') \\ &= \sum_{s_1 \rightarrow s_2 : x_A e \in R} \text{stoic}_r(A) x_A e. \end{aligned}$$

Or  $x_A^0 = 0$ , on a donc  $x_A = 0$ , donc les deux réseaux sont congruents.

Enfin, les règles (**CONTEXTE**) et (**SIMILARITÉ**) sont correctes, par induction sur l'arbre de preuve. En effet, si  $M \mid M'' \Rightarrow M' \mid M''$ , avec la règle (**CONTEXTE**) telle que  $M \Rightarrow M'$ , alors par induction  $M \sim M'$  et donc  $M \mid M'' \sim M' \mid M''$ . De même, si  $M_1 \Rightarrow M_2$  avec la règle (**SIMILARITÉ**), on a par le Lem. 8 :

$$M_1 \sim M'_1 \sim M'_2 \sim M_2.$$

□

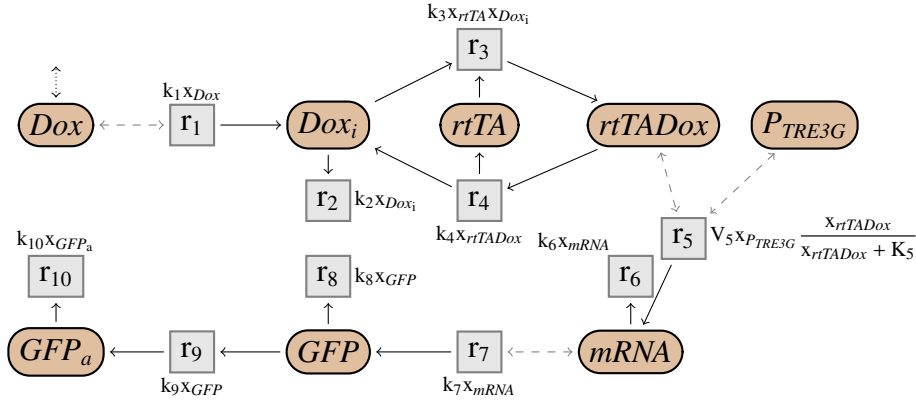


Fig 4.9: Graphe du réseau cinétique Tet-On.

### 4.3 Application au réseau cinétique Tet-On

On applique maintenant nos règles de simplification au réseau cinétique Tet-On (Gossen & Bujard, 1992; Gossen et al., 1995; Huang et al., 2010). On rappelle le graphe du réseau dans la Fig. 4.9.

Toutes les espèces, à part  $Dox$ , sont internes. On utilise la fonction d'observation  $\Omega_{\text{Tet-On}}$  de l'Ex. 11, qui n'observe que la quantité de  $GFP_a$ .

On peut tout d'abord utiliser la règle (**MODIFICATEUR CONSTANT**) pour supprimer le gène  $P_{TRE3G}$ , en remplaçant sa concentration par  $x_{P_{TRE3G}}^0$ . On obtient la réaction  $r'_5 = \emptyset \rightarrow mRNA$ ;  $V_5 x_{P_{TRE3G}}^0 \frac{x_{rtTADox}}{x_{rtTADox} + K_5}$ .

Ensuite, on utilise la règle (**INTERMÉDIAIRE - LINÉAIRE**) sur  $mRNA$ . Elle combine les réactions  $r'_5$  et  $r_6$ . On a  $T = V_5 x_{P_{TRE3G}}^0 \frac{x_{rtTADox}}{x_{rtTADox} + K_5}$  et  $T' = k_6$ . La fusion des deux réactions donne une réaction cyclique  $\emptyset \rightarrow \emptyset$ ;  $e$ , qui peut directement être supprimée avec la règle (**COMBINAISON**). On remplace  $x_{mRNA}$  par  $T/T'$  dans la réaction  $r_7$ , et on ajoute la contrainte  $x_{mRNA}^0 = \frac{V_5 x_{P_{TRE3G}}^0}{k_6} \frac{x_{rtTADox}}{x_{rtTADox} + K_5}$ . On obtient les réactions décrites dans la Fig. 4.10.

On peut procéder de la même façon pour l'espèce intermédiaire  $rtTADox$ . L'espèce  $rtTA$  devient alors un modificateur, et peut donc être supprimée avec la règle (**MODIFICATEUR CONSTANT**).

On utilise les règles de similarité afin d'obtenir des cinétiques qui dépendent de la quantité totale  $x_{rtTA}^0 + x_{rtTADox}^0$  de  $rtTA$ , liée ou non, comme dans l'Ex. 31 de Michaelis-Menten. On obtient les réactions décrites dans la Fig. 4.11, avec les



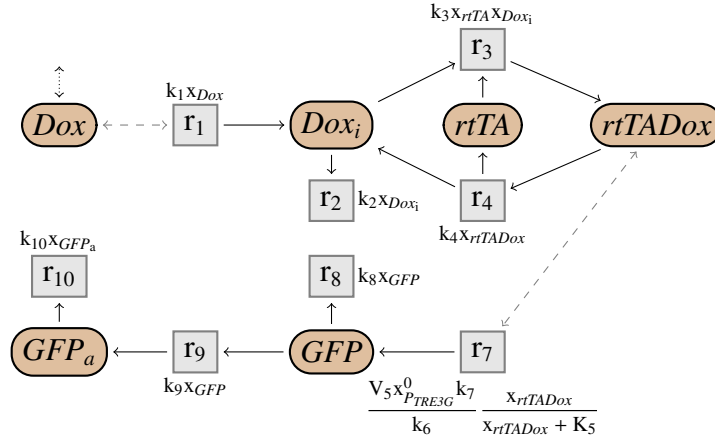


Fig 4.10: Graphe d'un réseau cinétique partiellement simplifié de Tet-On.

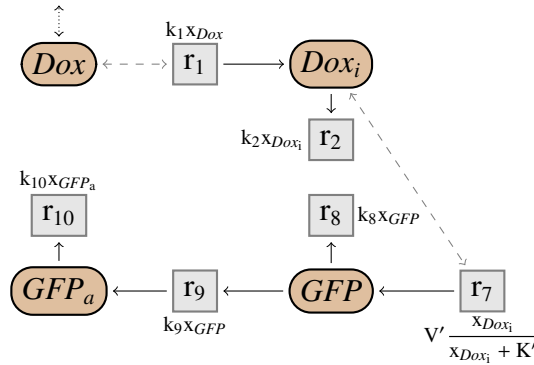


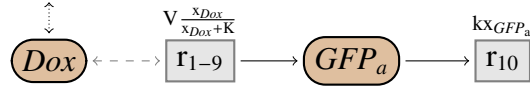
Fig 4.11: Graphe d'un réseau cinétique partiellement simplifié de Tet-On.

nouveaux paramètres suivants :

$$V' = \frac{k_7 V_5 x_{P_{TRE3G}}^0 (x_{rtTADox}^0 + x_{rtTA}^0)}{k_6 ((x_{rtTADox}^0 + x_{rtTA}^0) + K_5)} \quad K' = \frac{k_4 K_5}{k_3 ((x_{rtTADox}^0 + x_{rtTA}^0) + K_5)}$$

On peut alors utiliser la règle (**INTERMÉDIAIRE - LINÉAIRE**) sur les espèces  $Dox_i$  et  $GFP$ . On obtient le réseau simplifié  $M'_{Tet-On} = R' \&cst(x_{Dox}) \wedge C'$ . Ce réseau ne contient plus que deux espèces,  $Dox$  et  $GFP_a$ . Il n'a plus que deux réactions, illustrées dans la Fig. 4.12.

La contrainte  $C'$  représente la valeur des concentrations initiales des espèces



$$V = \frac{V_5 k_7 k_9 x_{P_{TRE3G}}^0 (x_{rtTA}^0 + x_{rtTADox}^0)}{k_6 (k_8 + k_9) (K_5 + (x_{rtTA}^0 + x_{rtTADox}^0))} \quad K = \frac{K_5 k_2 k_4}{k_1 k_3 (K_5 + (x_{rtTA}^0 + x_{rtTADox}^0))}$$

Fig 4.12: Graphe du réseau cinétique simplifié de Tet-On.

supprimées :

$$\begin{aligned} C' &= x_{GFP}^0 = \frac{k_4}{k_5} x_{mRNA}^0 \\ \wedge x_{mRNA}^0 &= \frac{V_5 x_{P_{TRE3G}}^0 x_{rtTADox}^0}{k_6 (x_{rtTADox}^0 + K_5)} \\ \wedge x_{rtTADox}^0 &= \frac{k_3}{k_4} x_{Dox_i}^0 x_{rtTA}^0 \\ \wedge x_{Dox_i}^0 &= k_1 / k_2 x_{Dox} \end{aligned}$$

Comme notre simplification est correcte, ce réseau est congruent au réseau initial.

## 4.4 Simplification de systèmes d'équations

Dans cette section, on transforme les règles de simplification de réseaux cinétiques vues dans ce chapitre afin d'obtenir des règles de simplification de systèmes d'équations. Les règles de simplification sont présentées dans la Fig. 4.13.

**Variable à l'équilibre** Commençons par considérer les règles de simplification d'espèces intermédiaires. Avec les axiomes de similarité présentés dans la Fig. 4.1, la règle (**VARIABLE STABLE**) d'élimination d'une variable  $x$  à l'équilibre consiste juste à remplacer  $\dot{x} = e$  par  $e = 0$ ,  $\dot{x} = 0$  et de remplacer  $x$  par  $x^0$ .

**Symétrie** Comme pour les réseaux de réactions cinétiques, on considère deux règles de symétrie entre variables. Dans la première règle, (**SYMÉTRIE**), les deux variables  $x_1$  et  $x_2$  ont la même *ODE*, et la même valeur initiale. Elles seront donc toujours égales, et on peut donc supprimer  $x_2$ , supposée non observable.

Dans la règle (**SYMÉTRIE - CONSOMMATION**), les variables  $x_1$  et  $x_2$  sont consommées de façon similaire, c'est-à-dire  $\dot{x}_1 = e_1 - x_1 e$  et  $\dot{x}_2 = e_2 - x_2 e$ . De plus, dans les autres équations,  $x_1$  et  $x_2$  apparaissent toujours sous la forme  $x_1 + x_2$ . On va alors remplacer  $x_1 + x_2$  par une nouvelle variable  $x$ , avec  $\dot{x} = (e_1 + e_2)[x_1 + x_2 := x] - x e$ .

**Variable Constante** Dans la règle (**VARIABLE CONSTANTE**), qui correspond à la règle (**MODIFICATEUR CONSTANT**), on substitue une variable  $x$  par sa valeur initiale  $x^0$  lorsqu'elle est constante au cours du temps. Contrairement à la règle (**VARIABLE STABLE**), le fait que  $x$  soit constante est une conséquence de  $\dot{x} = 0$ , et non pas d'une contrainte  $cst(x)$ .

**Variable Inutile** La règle (**ESPÈCE INUTILE**) peut se transformer en une règle (**VARIABLE INUTILE**), dans laquelle une variable  $x$ , non observable, n'apparaît que dans sa propre *ODE*.

**Variable Nulle** La règle (**RÉACTIONS INAPPLICABLES**) se transforme en une règle (**VARIABLE NULLE**), dans laquelle une variable  $x$  est nulle initialement et son *ODE* est de forme  $\dot{x} = x e$ , et donc  $x$  sera toujours nulle au cours du temps.

Toutes ces règles sont trivialement correctes.

**Proposition 12.** Soit  $E$  un système tel que  $E \Rightarrow E'$ , alors  $E \simeq E'$ .

(VARIABLE STABLE)

$$\frac{}{E \wedge \dot{x} = e \wedge cst(x) \Rightarrow E[x := x^0] \wedge \dot{x} = 0 \wedge 0 = e[x := x^0]}$$

(SYMÉTRIE)

$$\frac{\Omega(x_2) = \mathbb{O} \quad x_2 \notin Vars(e)}{E \wedge \dot{x}_1 = e \quad \wedge \dot{x}_2 = e[x_1 := x_2] \quad \wedge x_1^0 = x_2^0 \quad \Rightarrow \quad E[x_2 := x_1] \wedge \dot{x}_1 = e \quad \wedge x_1^0 = x_2^0}$$

(SYMÉTRIE - CONSOMMATION)

$$\frac{\Omega(x_1) = \Omega(x_2) = \Omega(x) = \mathbb{O} \quad x \notin Vars(E), Vars(e_1), Vars(e_2), Vars(e) \quad x_1, x_2 \notin Vars(e) \quad \forall e' \in Expr(E) \cup \{e_1, e_2\}. dep(e', x_1, x_1 + x_2) \wedge dep(e', x_2, x_1 + x_2)}{E \quad \wedge \dot{x}_1 = e_1 - x_1 e \quad \wedge \dot{x}_2 = e_2 - x_2 e \quad \Rightarrow \quad E[x_1 + x_2 := x] \quad \wedge \dot{x} = (e_1 + e_2)[x_1 + x_2 := x] - x e}$$

(VARIABLE CONSTANTE)

$$\frac{x \in Vars(E)}{E \wedge \dot{x} = 0 \Rightarrow E[x := x^0] \wedge \dot{x} = 0}$$

(VARIABLE INUTILE)

$$\frac{\Omega(x) = \mathbb{O} \quad x \notin Vars(E)}{E \wedge \dot{x} = e \Rightarrow E \wedge \dot{x} = 0}$$

(VARIABLE NULLE)

$$\frac{}{E \wedge \dot{x} = x e \wedge x^0 = 0 \Rightarrow E[x := 0] \wedge \dot{x} = 0 \wedge x^0 = 0}$$

(SIMILARITÉ)

$$\frac{E_1 \equiv E'_1 \quad E'_1 \Rightarrow E'_2 \quad E'_2 \equiv E_2}{E_1 \Rightarrow E_2}$$

Fig 4.13: Règles de simplification de systèmes d'équations, avec  $\Omega$  la fonction d'observation.

**Simulation** Il existe alors une notion de simulation entre les règles de simplification pour les réseaux cinétiques et celles pour les systèmes d'équations.

**Proposition 13.** Soit  $M$  un réseau tel que  $M \Rightarrow M'$ , alors  $E(M) \Rightarrow^* E(M')$ .

*Proof.* On considère tout d'abord la règle de contexte (**CONTEXTE**). Supposons que  $M \Rightarrow M'$ . Alors il existe toujours des sous-réseaux  $M_1$ ,  $M'_1$  et  $M''$  tels que  $M = M_1 \mid M''$  et  $M' = M'_1 \mid M''$ , et tels que  $M_1 \Rightarrow M'_1$  avec une règle différente de (**CONTEXTE**). On prouve alors dans la suite que quelque soit cette règle, on a bien  $E(M) \Rightarrow E(M')$ . On note  $R''$  et  $C''$  les réactions et la contrainte de  $M''$ .

Si  $M_1 \Rightarrow M'_1$  en utilisant une des règles de la Fig. 4.3 sur une espèce  $A$ , alors on a  $cst(x_A)$  dans le réseau initial, ainsi que dans son système d'équations  $E(M)$ . On peut donc appliquer la règle (**VARIABLE STABLE**), et obtenir un système  $E'$ . On va prouver que  $E' \equiv E(M')$ . Notons que par définition, on a  $A \notin \text{Espèces}(M'')$ .

Trivialement, l'*ODE* pour  $x_A$  est la même dans les deux systèmes :  $\dot{x}_A = 0$ . On considère tout d'abord le cas de (**INTERMÉDIAIRE - LINÉAIRE**). Dans  $E(M')$ , on a une contrainte  $x_A^0 = F$ . Dans  $E'$ , on a la contrainte  $0 = T - x_A^0 T'$ , qui est bien similaire à  $x_A^0 = F$ . Soit  $B \neq A$  une autre espèce. Son *ODE* dans  $E'$  est :

$$\dot{x}_B = \sum_{r \in M} \text{stoic}_r(B) \text{kin}(r)[x_A := x_A^0]$$

Dans  $E(M')$ , son équation différentielle peut se réécrire de la façon suivante (on ne détaille pas tous les calculs, qui sont similaires à ceux vus dans la preuve du Théorème 4) :

$$\begin{aligned} \dot{x}_B &= \sum_{r=s_1 \rightarrow s_2+aA; e \in R_{\text{prod}}; r'=s'_1+a'A \rightarrow s'_2; x_A e' \in R_{\text{cons}}} \sum (a' s_2(B) + a s'_2(B) - a' s_1(B) - a s'_1(B)) e e' / T' \\ &\quad + \sum_{r \in R_{\text{mod}} \cup R''} \text{stoic}_r(B) \text{kin}(r)[x_A := F] \\ &\equiv \sum_{r=s_1 \rightarrow s_2+aA; e \in R_{\text{prod}}} \text{stoic}_r(B) e + \sum_{r'=s'_1+a'A \rightarrow s'_2; x_A e' \in R_{\text{cons}}} \text{stoic}_{r'}(B) e' F \\ &\quad + \sum_{r \in R_{\text{mod}} \cup R''} \text{stoic}_r(B) \text{kin}(r)[x_A := x_A^0] \\ &\equiv \sum_{r \in R_{\text{prod}}} \text{stoic}_r(B) \text{kin}(r) + \sum_{r' \in R_{\text{cons}}} \text{stoic}_{r'}(B) \text{kin}(r')[x_A := x_A^0] \\ &\quad + \sum_{r \in R_{\text{mod}} \cup R''} \text{stoic}_r(B) \text{kin}(r)[x_A := x_A^0] \\ &= \sum_{r \in M} \text{stoic}_r(B) \text{kin}(r)[x_A := x_A^0]. \end{aligned}$$

Pour la contrainte  $C'$  (dans  $M_1$ ), elle est remplacée par  $C'[x_A := x_A^0]$  dans  $E'$ , et par  $C'[x_A := F]$ . Ces deux contraintes sont bien similaires puisque dans les

deux systèmes on a aussi  $x_A^0 = F$ . La contrainte  $C''$  n'est pas modifiée, puisque  $x_A \notin \text{Vars}(C'')$ .

Les règles (INTERMÉDIAIRE - INVERSIBLE) et (INTERMÉDIAIRE - DÉPENDANTE) fonctionnent de la même manière.

Pour la règle (COMBINAISON), on a directement  $E(M) = E(M')$ .

Pour la règle (SYMÉTRIE) sur  $A_1$  et  $A_2$ , on peut directement l'imiter avec la règle (SYMÉTRIE) sur  $x_{A_1}$  et  $x_{A_2}$ . On obtient trivialement le même système. La règle (SYMÉTRIE - RÉACTANT) peut quant à elle être imitée par (SYMÉTRIE - CONSOMMATION).

Pour la règle (FACTORISATION), on a directement  $E(M) = E(M')$ .

La règle (MODIFICATEUR CONSTANT) est imitable trivialement par la règle (VARIABLE CONSTANTE), la règle (ESPÈCE INUTILE) par (VARIABLE INUTILE), tandis que la règle (RÉACTIONS INAPPLICABLES) est imitable avec (VARIABLE NULLE). La règle de similarité est imitée par la règle correspondante pour les systèmes.  $\square$

## 4.5 Discussion

Dans cette section, on discute et compare notre simplification avec d'autres simplifications dans la littérature. On étudie tout d'abord les différentes conditions d'équilibre qui existent dans la littérature permettant de simplifier des modèles biologiques, en préservant leur comportement, en général de façon approchée. On discute ensuite de l'importance de la structure des réseaux de réactions, en particulier pour la simplification. On étudie ensuite les méthodes de simplification en plusieurs étapes, séparant le calcul de la structure simplifiée, à l'aide par exemple des modes élémentaires, et celui des cinétiques. Enfin, on discute des différentes conditions sur les cinétiques dans les règles de simplification et de leurs applications en pratique.

### 4.5.1 Conditions d'équilibre

La plupart des simplifications de modèles biologiques sont basées sur l'existence de *différents ordres de grandeurs* entre les vitesses des réactions. En se plaçant sur un certain intervalle de temps, on peut alors supposer que les réactions très rapides permettent d'atteindre rapidement un équilibre, tandis que les réactions très lentes n'ont qu'un effet négligeable sur le comportement global du système. C'est par exemple le cas des méthodes basées sur les variétés de faibles dimensions du point de vue géométrique (voir par exemple (Fraser, 1988), (Roussel & Fraser, 1990), (Maas & Pope, 1992), (Stiefenhofer, 1998), (Davis & Skodje, 1999) ou encore (Gorban & Karlin, 2003)), ou de celles basées sur les perturbations singulières (CSP) en théorie du contrôle (par exemple les travaux de (Heineken, Tsuchiya, & Aris, 1967), (Lam & Goussis, 1994), (Lu, Ju, & Law, 2001) ou (Zagaris, Kaper, & Kaper, 2004)). C'est aussi le cas des méthodes basées sur la quasi-stabilité (QSSA) d'espèces chimiques (comme par exemple (Bodenstein, 1913), (Segel & Slemrod, 1989) ou encore (Turanyi, Tomlin, & Pilling, 1993)), sur le quasi-équilibre (QE) de flux de réactions (comme dans (Gorban, Karlin, Ilg, & Öttinger, 2001) ou (Gibbs, 2014)) ou encore sur les équilibres tropicaux entre les flux dominants (Radulescu, Vakulenko, & Grigoriev, 2015).

Ces simplifications sont néanmoins en général des *approximations*. Au contraire, notre méthode de simplification cherche à préserver *exactement* le comportement déterministe. Il est alors nécessaire d'avoir des conditions exactes d'équilibre pour les espèces intermédiaires, comme dans les règles présentées dans la Fig. 4.3.

Considérons par exemple le cas de la classique simplification enzymatique proposée par Michaelis et Menten (Michaelis & Menten, 1913), dont les réactions du réseau initial sont illustrées dans la Fig. 2.19. C'est en général une approximation, et de nombreuses études se sont intéressées aux différentes conditions

pour qu'elle soit valable (par exemple les travaux de (Palsson, 1987), (Segel, 1988), (Segel & Slemrod, 1989), (Schnell & Mendoza, 1997) ou encore (Sanft et al., 2011)). La simplification n'est possible que si le système vérifie un quasi-équilibre entre les réactions  $r_1$  et  $r_2$  (Michaelis & Menten, 1913), ou une quasi-stabilité (Briggs & Haldane, 1925) du complexe  $C$ . Pour notre simplification, on suppose que la stabilité de  $C$  est exacte. Remarquons qu'en général, pour la quasi-stabilité, on suppose également que la concentration initiale du substrat  $S$  est très supérieure à celle de l'enzyme (et du complexe). Une conséquence est donc que la quantité de  $S$  ne va être modifiée que de façon négligeable, en d'autres termes, que  $S$  est également quasi-stable. Dans notre cas, il est également nécessaire que  $S$  soit à l'équilibre pour que  $C$  le soit et qu'on puisse simplifier le modèle.

Une autre approche consiste à supposer non pas qu'une espèce  $A$  est dans un état d'équilibre, mais plutôt qu'une solution, par exemple  $A + 2B$ , est à l'équilibre, c'est-à-dire  $\dot{x}_A + 2\dot{x}_B = 0$ . Elle se base sur le concept des graphes de solutions (souvent appelé *graphe des complexes*), introduit dans (Feinberg, 1987; Horn, 1972; Horn & Jackson, 1972), de façon similaire aux réseaux de Kron dans les réseaux électriques (Kron, 1939). Cette approche a notamment été utilisée dans les travaux de (Gunawardena, 2012) pour calculer un système simplifié d'équations, ou dans (S. Rao, Van der Schaft, Van Eunen, Bakker, & Jayawardhana, 2014) pour simplifier (à l'aide d'approximations) des réseaux de réactions. Cependant, la simplification de solutions plutôt que d'espèces peut entraîner la disparition de liens et de dépendances entre espèces, comme expliqué dans (S. Rao, Jayawardhana, & van der Schaft, 2016).

### 4.5.2 Importance de la structure

Comme discuté dans l'introduction, préserver la structure des réseaux de réactions est plus avantageux que de connaître uniquement les équations. En effet, un réseau peut être représenté graphiquement et est plus facile à comprendre. De plus, il permet de dériver les sémantiques non déterministe, déterministe et stochastique, et d'adopter une approche contextuelle.

La structure peut également être utile pour simplifier un modèle biologique. Par exemple, King et Altmann (King & Altman, 1956) cherchent à calculer la vitesse de production d'un produit dans une chaîne de réactions enzymatiques, en supposant que les complexes sont à l'équilibre. Ils utilisent une méthode graphique, basée sur la présence de motifs dans la structure du réseau. Leur méthode a été améliorée, afin d'éviter au maximum des calculs inutiles, et a été généralisée en utilisant des méthodes de la théorie des graphes (voir par exemple (Wong & Hanes, 1962), (Volkenstein & Goldstein, 1966), (Cha, 1968) ou (Cornish-Bowden, 1977)).

Le même type de méthode a alors été utilisée pour simplifier des espèces inter-



médiaires de façon plus générale, sans se limiter aux cas enzymatiques, et calculer des systèmes d'équations réduits (Kuo-Chen & Forsen, 1981; Chou, 1981; Zhou & Deng, 1984; Feliu & Wiuf, 2012). Dans (Sáez, Wiuf, & Feliu, 2016), une méthode graphique, basée sur le calcul d'arbres couvrants dans les graphes des espèces intermédiaires, permet de les simplifier, cette fois dans des réseaux de réactions. On étudiera plus précisément le lien entre cette simplification et la notre dans la Sec. 5.4 du chapitre suivant.

La structure permet également de pouvoir définir des notions de *bisimulations* dans les réseaux de réactions. Par exemple, dans (Cardelli, Tribastone, Tschaikowski, & Vandin, 2015), les auteurs proposent deux notions de bisimulation dans les réseaux avec une sémantique déterministe. La *bisimulation vers l'avant* permet de regrouper des espèces qui sont consommées de la même façon, de façon similaire à la règle (SYMÉTRIE - RÉACTANT). La *bisimulation vers l'arrière* correspond au cas d'espèces ayant le même comportement, comme dans la règle (SYMÉTRIE). Les mêmes notions de bisimulation ont également été définies pour les systèmes d'équations différentielles dans (Cardelli, Tribastone, Tschaikowski, & Vandin, 2016).

Dans (Cardelli, 2014), les auteurs proposent également plusieurs notions de morphismes entre deux réseaux, permettant de fusionner des espèces et des réactions. Ils utilisent alors des propriétés sur la structure des réseaux pour obtenir des résultats sur leurs sémantiques déterministes. Notamment, le morphisme doit préserver la stœchiométrie et les réactants des réactions. Ces propriétés ne sont néanmoins pas vérifiées en général avec notre simplification. Par exemple, le morphisme correspondant à notre règle (SYMÉTRIE - RÉACTANT), qui fusionne deux espèces symétriques, préserve les réactants mais pas la stœchiométrie, avec les notions définies dans (Cardelli, 2014).

### 4.5.3 Modes élémentaires

Un autre type de méthodes de simplification de réseau avec une sémantique déterministe consiste à calculer séparément la structure simplifiée et les cinétiques. Le calcul de la structure peut alors se faire de façon similaire aux *modes élémentaires* (Schuster & Hilgetag, 1994).

Un *mode*, aussi appelé parfois *chemin de réactions* (comme par exemple dans (Horiuti & Nakamura, 1957)), est un multi-ensemble de réactions qui, si elles sont appliquées ensembles, ne modifient pas le nombre de molécules des espèces intermédiaires. À chaque mode correspond alors une réaction globale, simplifiée, sans espèce intermédiaire. Il existe alors un (unique) ensemble de modes minimaux, les *modes élémentaires*, permettant de générer l'ensemble de tous les modes du réseau par combinaison linéaire. Géométriquement, l'ensemble des modes forme un *cône*, et les modes élémentaires correspondent alors aux *arrêtes extrêmes* de ce

cône (Gagneur & Klamt, 2004). Ce sont également les *flux extrêmes* en analyse stœchiométrique des réseaux (voir (Clarke, 1980)), et les modes sont des *invariants de transitions* dans les réseaux de Pétri (voir par exemple (Zevedei-Oancea & Schuster, 2003)).

Clarke (Clarke, 1992), reprenant et améliorant les résultats proposés par Masuda (Masuda, 1990), propose alors de calculer indépendamment les réactions globales correspondant aux modes élémentaires d'une part, et d'autre part un système simplifié d'équations, puis d'associer à chaque réaction un taux cinétique afin de correspondre aux équations simplifiées. Dans (Radulescu et al., 2008), Radulescu et al. utilise la même méthode, en calculant un système simplifié d'équations, à l'aide d'approximations basées sur les cinétiques dominantes ou les équilibres tropicaux (Radulescu et al., 2015).

La difficulté est alors de déterminer les taux cinétiques des nouvelles réactions correspondants au système simplifié. Si cela peut parfois être fait de façon simple et unique, ce n'est pas le cas pour des systèmes plus complexes, notamment lorsque certaines réactions simplifiées sont la combinaison d'autres (comme dans la règle (COMBINAISON)). Dans (Radulescu, Gorban, Zinovyev, & Noel, 2012), Radulescu et al. propose alors d'éliminer ces réactions avant de calculer les taux cinétiques. On étudiera plus en détail le lien entre la simplification proposée dans (Radulescu et al., 2008) et la notre, qui calcule dans le même temps la structure des réactions et leurs cinétiques, dans le chapitre suivant.

#### 4.5.4 Conditions d'application

Afin de conserver le comportement déterministe de façon exacte, certaines règles de simplification nécessitent des conditions fortes sur les cinétiques. Dans les règles de symétrie présentées dans la Fig. 4.6, les paramètres des réactions impliquant les espèces symétriques doivent être les mêmes. Ces règles ne peuvent par exemple pas être utilisées pour simplifier le système MAPK.

Les règles d'élimination des espèces intermédiaires présentées dans la Fig. 4.3 nécessitent également des conditions particulières sur les cinétiques des réactions qui consomment les espèces intermédiaires. Néanmoins, contrairement aux simplifications exactes de réseaux existantes dans la littérature (voir par exemple (Sáez et al., 2016)), nos règles ne sont pas limitées au cas linéaire.

## 4.6 Conclusion

Dans ce chapitre, on a présenté un ensemble de règles de simplification, préservant la sémantique déterministe. Ces règles sont basées sur les réseaux de réactions et non les systèmes d'équations. Elles calculent en même temps la structure des réactions et leurs cinétiques simplifiées. Elles sont symboliques, et contextuelles. Ces règles permettent notamment d'éliminer des espèces intermédiaires, lorsqu'elles sont à l'équilibre. Les conditions sur les cinétiques permettent de traiter certains cas non linéaires. Un autre ensemble de règles permet de fusionner des espèces se comportant de façon symétrique. D'autres règles peuvent simplifier des réactions, combiner d'autres réactions, factoriser des réactions, ou encore supprimer des espèces inutiles. Ces règles peuvent être modifiées pour simplifier les systèmes d'équations au lieu des réseaux.

Comme pour le cas non déterministe, la simplification n'est pas complète, et on peut donc trouver et ajouter de nouvelles règles. La simplification pourrait être utilisée de façon systématique sur les modèles de la base [Biomodels](#), afin de les simplifier et les comparer. Il serait intéressant également de voir s'il est possible de relâcher un peu les conditions d'application des règles. Dans ce cas, le réseau obtenu serait une approximation, et il faudrait alors estimer et borner formellement l'erreur introduite par la simplification.

## CHAPITRE 5

---

### Confluence de l'élimination des espèces intermédiaires

---

Dans ce chapitre, on étudie la confluence de l'élimination des espèces intermédiaires dans les réseaux de réactions cinétiques. Comme on va le voir, les règles de simplification présentées dans la Fig. 4.3 du Chapitre 4 ne sont pas confluentes. Il y a ici trois sources différentes de non confluence : les *réactions dépendantes*, la *non-préservation* des conditions de simplification, et la *non confluence des cinétiques*.

Les réactions dépendantes sont des réactions pouvant être introduites par la simplification des espèces intermédiaires du réseau, et qui sont la combinaison d'autres réactions du réseau. Selon l'ordre d'élimination des espèces intermédiaires, ces réactions peuvent être présentes ou non. Les réseaux simplifiés n'auront alors pas les mêmes structures, c'est-à-dire pas les mêmes réactions, même en négligeant les cinétiques.

On étudie tout d'abord ce problème en négligeant les cinétiques. On montre que pour pouvoir supprimer ces réactions de façon confluyente, il est nécessaire d'introduire une notion de *réseaux de flux*, où chaque flux dans les réseaux simplifiés est une combinaison linéaire des réactions du réseau initial. On peut alors ajouter une nouvelle règle pour supprimer les flux dépendants, et obtenir une simplification confluyente (sans les cinétiques). On montre également qu'il y a un lien entre notre simplification et le calcul des modes élémentaires d'un réseau.

Le deuxième problème pour la confluence est que les conditions d'application

des règles, comme par exemple la linéarité des expressions cinétiques, ne sont pas préservées par la simplification. Par exemple, si deux espèces intermédiaires  $A$  et  $B$  sont linéaires dans un réseau  $M$ , et qu'on le simplifie en  $M'$  en supprimant l'espèce  $A$  avec la règle (**INTERMÉDIAIRE - LINÉAIRE**), alors il est possible que l'espèce  $B$  ne soit plus linéaire dans  $M'$ , et donc ne puisse plus être supprimée.

En se restreignant au cas d'espèces intermédiaires linéaires, on montre qu'en ajoutant des conditions supplémentaires sur les réseaux, et notamment sur la stœchiométrie des espèces intermédiaires, on est sûr de pouvoir toujours simplifier l'ensemble des espèces intermédiaires et des flux dépendants. Les réseaux complètement simplifiés n'auront donc plus aucune espèce intermédiaire, et auront tous la même structure.

Néanmoins, les expressions cinétiques des flux des réseaux simplifiés peuvent toujours être différentes. En effet, l'élimination d'un flux dépendant peut parfois se faire de plusieurs manières, produisant la même structure mais pas les mêmes expressions cinétiques. On illustre notamment ce cas sur un exemple biologique, où la simplification d'espèces intermédiaires peut produire deux réseaux différents, avec différentes cinétiques.

On donne alors le critère suivant sur la structure du réseau initial : un réseau est décomposable de façon unique si tout mode du réseau a une unique décomposition en modes élémentaires. On prouve alors que lorsque ce critère est vérifié, alors la confluence est totale, c'est-à-dire à la fois sur la structure et sur les expressions cinétiques.

Enfin, on compare cette simplification à la littérature, en particulier aux simplifications de (Radulescu et al., 2008) et de (Sáez et al., 2016).

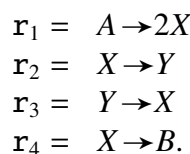
Ce chapitre reprend les résultats présentés dans (Madelaine, Tonello, et al., 2016a) et dans (Madelaine, Tonello, et al., 2016b).

## 5.1 Confluence de la simplification sans cinétique

On étudie le problème de la confluence de l'élimination des espèces intermédiaires dans le cas sans cinétique. On illustre tout d'abord le problème de confluence posé par l'élimination des espèces intermédiaires sur un exemple. On introduit alors une représentation d'un réseau de réactions sous la forme de vecteurs, par rapport à une base de réactions initiales. On montre qu'avec cette représentation, on peut ajouter de nouvelles règles et obtenir une simplification confluyente (toujours sans cinétique). On montre également que notre algorithme de simplification peut alors être utilisé pour calculer les modes élémentaires d'un réseau.

### 5.1.1 Non confluence de l'élimination des espèces intermédiaires

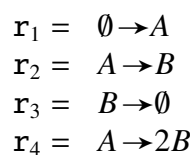
On considère l'ensemble de réactions (sans cinétique)  $R$  suivant, illustré dans la Fig. 5.1 :



On veut supprimer les deux espèces intermédiaires  $X$  et  $Y$ . En commençant par l'espèce  $X$ , avec la règle (**INTERMÉDIAIRE - LINÉAIRE**) du Chapitre 4, on obtient les réactions  $R_X$ , illustrées dans la Fig. 5.1. Par exemple, la réaction  $r_{12^2}$  représente la combinaison de la réaction  $r_1$  avec deux fois la réaction  $r_2$ , et transforme donc une molécule de  $A$  en deux molécules de  $Y$ . On peut alors supprimer  $Y$ , et obtenir les réactions  $R_{XY}$ . Inversement, on peut commencer par supprimer  $Y$ , puis  $X$ , et obtenir les réactions  $R_{YX}$ .

On remarque alors que  $R_{XY}$  possède une réaction supplémentaire, la réaction  $r_{12^23^24^2}$ . Donc, si on ne considère que les règles d'élimination d'espèces intermédiaires, la simplification n'est pas confluyente. Néanmoins, cette réaction est en fait la combinaison d'autres réactions. On pourrait par exemple la supprimer avec la règle (**COMBINAISON**) présentée dans le chapitre précédent. Malheureusement, comme l'illustre l'exemple suivant, la règle (**COMBINAISON**) n'est pas non plus confluyente.

**Exemple 41.** On considère les réactions suivantes :



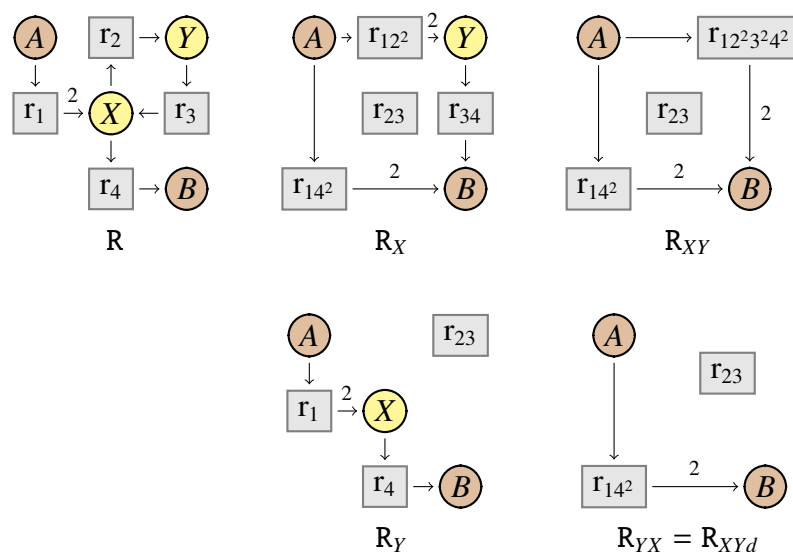


Fig 5.1: Élimination des espèces intermédiaires  $X$  et  $Y$  dans les réactions  $R$ , dans différents ordres, illustrant le besoin d'élimination des réactions dépendantes.

La réaction  $r_4$  est la combinaison de la réaction  $r_1$  et deux fois la réaction  $r_2$ , et on peut donc la supprimer avec la règle (**COMBINAISON**) (en négligeant les cinétiques), et obtenir les réactions  $\{r_1, r_2, r_3\}$ . Mais on peut aussi supprimer la réaction  $r_2$ , qui est la combinaison de  $r_3$  et  $r_4$ , et obtenir  $\{r_1, r_3, r_4\}$ . On ne peut alors plus simplifier ni les réactions  $\{r_1, r_2, r_3\}$  ni  $\{r_1, r_3, r_4\}$ , donc la simplification n'est pas confluente.

Le problème ici vient du fait qu'une réaction  $r$  peut être la combinaison d'un ensemble de réactions contenant une réaction  $r'$ , et qu'en même temps  $r'$  peut être la combinaison de réactions contenant  $r$ . On propose dans la suite une autre notion de combinaison de réactions, permettant d'éviter ce cas.

Revenons aux réactions de  $R_{XY}$ . On veut supprimer la réaction  $r_{12^2 3^2 4^2}$ , mais avec une règle confluente. Au lieu d'utiliser la règle (**COMBINAISON**), qui utilise les vecteurs de stœchiométrie, on va se servir des liens avec les réactions initiales. En effet, par construction, on peut voir que la réaction  $r_{12^2 3^2 4^2}$  est la combinaison d'une fois la réaction  $r_1$ , et de deux fois les réactions  $r_2$ ,  $r_3$  et  $r_4$ . On peut alors représenter cette réaction comme un vecteur  $u_{12^2 3^2 4^2} = (1, 2, 2, 2)$ , dans la base composée des réactions initiales ( $r_1, r_2, r_3, r_4$ ). On peut procéder de même pour les deux autres réactions, et obtenir les vecteurs  $u_{14^2} = (1, 0, 0, 2)$  et  $u_{23} = (0, 1, 1, 0)$ . On remarque alors que  $u_{12^2 3^2 4^2} = u_{14^2} + 2u_{23}$ . On dit alors que la réaction  $r_{12^2 3^2 4^2}$  est *dépendante* des deux autres réactions, et on peut alors la supprimer. On montrera plus loin qu'on ne peut jamais avoir une réaction  $r$  qui

dépend de  $\mathbf{r}'$  et en même temps que  $\mathbf{r}'$  dépende de  $\mathbf{r}$ .

### 5.1.2 Réseaux de flux

Dans cette section, on formalise les notions précédentes : les vecteurs et réseaux de flux, les réactions dépendantes et les nouvelles règles de simplification.

Soit  $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_m) \in \text{Réactions}^m$  un  $m$ -uplet fixé de réactions sans cinétique. Un vecteur  $u = (u_1, \dots, u_m) \in \mathbb{Z}^m$ , appelé *flux*, peut alors être utilisé pour représenter la réaction  $u\mathbf{r}$ , représentant  $u_1$  fois la réaction  $\mathbf{r}_1$ , combinée avec  $u_2$  fois la réaction  $\mathbf{r}_2$ , etc. Formellement,  $u\mathbf{r}$  est la réaction normalisée vérifiant, pour toute espèce  $A$  :

$$\text{stoic}_{u\mathbf{r}}(A) = \sum_{1 \leq i \leq m} u_i \text{stoic}_{\mathbf{r}_i}(A).$$

Un *réseau de flux*  $U$  est alors tout simplement un ensemble de flux, et représente l'ensemble de réactions  $U\mathbf{r} = \{u\mathbf{r} \mid u \in U\}$ .

### 5.1.3 Simplification

On veut alors simplifier, de façon confluente, les réseaux de flux en supprimant les espèces intermédiaires. La combinaison d'un flux produisant du  $X$  et d'un flux en consommant peut alors être simplement définie comme la somme des flux, pondérée par la stœchiométrie de  $X$  :

$$u \diamond_X u' = -\text{stoic}_{u'\mathbf{r}}(X)u + \text{stoic}_{u\mathbf{r}}(X)u'.$$

On a trivialement  $\text{stoic}_{(u \diamond_X u')\mathbf{r}}(X) = 0$ , donc la combinaison des réactions élimine bien l'espèce intermédiaire. On redéfinit alors dans la Fig. 5.2 la règle (**INTERMÉDIAIRE**) pour les réseaux de flux.

On dit qu'un flux  $u \in U$  dépend des flux  $u_1, \dots, u_k \in U \setminus \{u\}$  avec les coefficients  $a_1, \dots, a_k \in \mathbb{N}_{>0}$  si :

$$u = \sum_{1 \leq i \leq k} a_i u_i.$$

La règle (**DÉPENDANT**) supprime un flux dépendant. On ne peut jamais avoir un vecteur  $u$  qui dépend de  $u'$  et en même temps  $u'$  qui dépend de  $u$ . En effet, supposons que ce soit le cas, on aurait donc  $u = a'u' + \sum_i a_i u_i$  et  $u' = au + \sum_j a'_j u_j$ ,

d'où  $u = aa'u + \sum_i a_i u_i + \sum_j a' a'_j u_j$ , avec  $aa' > 0$ . Cela implique  $u_i = u_j = 0$  pour tout  $i, j$ , et  $a = a' = 1$ , d'où  $u = u'$ , ce qui contredit la définition de dépendance. Donc, contrairement au cas de la règle (**COMBINAISON**), la suppression des réactions dépendantes peut se faire de façon confluente.



$$\begin{array}{c}
\text{(INTERMÉDIAIRE)} \\
\hline
U \Rightarrow^{\text{Inter}} \frac{X \in \mathcal{I}}{\{u \diamond_X u' \mid u, u' \in U, \text{stoic}_{u\mathbf{r}}(X) > 0, \text{stoic}_{u'\mathbf{r}}(X) < 0\} \cup \{u \in U \mid \text{stoic}_{u\mathbf{r}}(X) = 0\}} \\
\\
\text{(DÉPENDANT)} \\
\hline
U \cup \{\sum_{1 \leq i \leq k} a_i u_i\} \Rightarrow^{\text{Dep}} U \quad \frac{u_1, \dots, u_k \in U \quad a_1, \dots, a_k \in \mathbb{N}_{>0}}{} \\
\\
\text{(FACTORISATION)} \\
\hline
U \cup \{au\} \Rightarrow^{\text{Fact}} U \cup \{u\} \quad \frac{a \in \mathbb{N}, a > 1}{}
\end{array}$$

Fig 5.2: Règles de simplification des réseaux de flux sans cinétique, avec  $\mathcal{I}$  l'ensemble des espèces internes.

**Exemple 42.** Soit  $\mathbf{r} = \{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$  les réactions du réseau  $N$  de la Fig. 5.1. Soit  $U = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$  le réseau de flux tel que  $U\mathbf{r} = N$ .

En simplifiant les espèces intermédiaires  $X$  puis  $Y$  dans  $U$ , on obtient le réseau de flux  $U_{XY} = \{(1, 0, 0, 2), (0, 1, 1, 0), (1, 2, 2, 2)\}$  tel que  $U_{XY}\mathbf{r} = N_{XY}$ . On note  $u_{14^2}$  le flux  $(1, 0, 0, 2)$ ,  $u_{23}$  le flux  $(0, 1, 1, 0)$ , etc.

On peut alors utiliser la règle (DÉPENDANT) sur le flux  $u_{12^23^24^2}$ , qui dépend de une fois  $u_{14^2}$  et de deux fois  $u_{23}$ , et obtenir le réseau  $U_{XYd}$  tel que  $U_{XYd}\mathbf{r} = N_{XYd} = N_{YX}$ .

On présente également la règle (FACTORISATION) adaptée aux réseaux de flux. Elle remplace un flux de forme  $au$  par le flux  $u$ . Elle est nécessaire pour être confluent, comme l'illustre l'exemple suivant.

**Exemple 43.** On considère le réseau  $U$  représenté dans la Fig. 5.3, avec trois espèces intermédiaires  $X$ ,  $Y$  et  $Z$ . Selon l'ordre de simplification des espèces intermédiaires, on peut obtenir soit le réseau  $U_{XYZ}$ , soit le réseau  $U_{YXZ}$ . Pour être confluent, il faut donc la règle (FACTORISATION), qui simplifie  $U_{XYZ}$  en  $U_{XYZf} = U_{YXZ}$ .

Notons qu'on ne définit pas ici de règle pour le contexte. En effet, ce n'est pas nécessaire, car les règles ci-dessus sont déjà définies de façon globales : si on peut simplifier  $U$  en  $U'$  avec l'une de ces règles, alors pour tout  $U''$  tel que  $\text{Espèces}(U'') \cap \mathcal{I} = \emptyset$ , on peut directement simplifier  $U \cup U''$  en  $U' \cup U''$  avec la même règle. Cela permet notamment de simplifier les preuves de confluence.

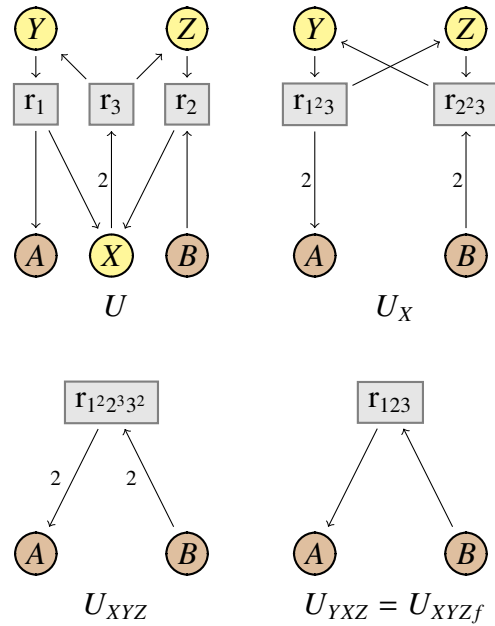


Fig 5.3: Élimination des espèces intermédiaires  $X$ ,  $Y$  et  $Z$  dans le réseau  $U$ , dans deux ordres différents, illustrant le besoin de la règle (FACTORISATION).

Remarquons que bien qu'on ait ici une simplification sans cinétique, il n'y a pas de lien direct avec la sémantique des attracteurs présentée dans le Chapitre 3. En effet, l'élimination des espèces intermédiaires est ici basée sur l'existence d'un équilibre (avec les cinétiques), tandis que celle pour la sémantique des attracteurs repose sur le fait que certaines réactions ne sont pas observables.

#### 5.1.4 Modes élémentaires

On va montrer que les trois règles de simplification ci-dessus permettent de calculer les *modes élémentaires* d'un réseau.

Soit  $\mathbf{r} = \{r_1, \dots, r_m\}$  la base fixée de  $m$  réactions, et  $\text{Espèces} = \{A_1, \dots, A_n\}$  l'ensemble (ordonné) des espèces moléculaires. On note  $S$  sa matrice de stœchiométrie de taille  $n \times m$ , telle que l'élément sur la  $j^{\text{ème}}$  ligne et la  $i^{\text{ème}}$  colonne soit égal à  $\text{stoc}_{r_i}(A_j)$ . Par exemple, la matrice de stœchiométrie  $S$  des réactions  $R$  de la Fig. 5.1 est représentée à gauche de la Fig. 5.4.

Soit  $\mathcal{J} \subseteq \mathcal{I}$  un ensemble d'espèces intermédiaires que l'on souhaite supprimer, et  $S_{\mathcal{J}}$  la projection de  $S$  sur les espèces  $\mathcal{J}$ . Pour tout vecteur  $u = (u_1, \dots, u_m)$ , on définit son *support* comme  $\text{supp}(u) = \{i \in \{1, \dots, m\} \mid u_i \neq 0\}$ .

On veut supprimer les espèces intermédiaires lorsqu'elles sont à l'équilibre.

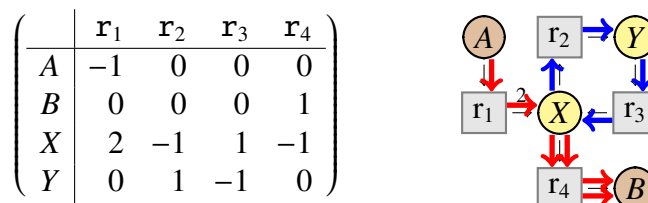


Fig 5.4: Matrice de stœchiométrie et graphe des réactions R, avec en rouge le mode élémentaire  $x_1 = (1, 0, 0, 2)$  et en bleu le mode élémentaire  $x_2 = (0, 1, 1, 0)$ .

Un *mode* est alors un flux, une combinaison de réactions, qui ne modifie pas la quantité de molécules des espèces intermédiaires. Un mode est *élémentaire* s'il est de taille minimale et n'est pas la combinaison d'autres modes.

**Définition 20** (Mode élémentaire). Un *mode élémentaire* pour une matrice  $S_{\mathcal{J}}$  sur  $\mathbb{N}$  de taille  $n \times m$  est un vecteur  $u \in \mathbb{N}^m$  tel que :

- $u$  est un mode :  $S_{\mathcal{J}}u = 0$ ,
- $u$  n'est pas combinaison d'autres modes : il n'existe pas de mode  $u'$  tel que  $S_{\mathcal{J}}u' = 0$ , avec  $\text{supp}(u) \subsetneq \text{supp}(u')$ ,
- $u$  est de taille minimale : il n'existe pas de mode  $u'$  tel que  $S_{\mathcal{J}}u' = 0$ , avec  $u = au'$  pour un entier naturel  $a > 1$ .

Tout mode est alors une combinaison linéaire positive (avec des coefficients entiers) des modes élémentaires (voir par exemple (Schuster & Hilgetag, 1994)).

Les modes élémentaires sont également liés au comportement stable d'un réseau. Supposons que toutes les espèces intermédiaires  $\mathcal{J}$  soient stables, et soit  $x = (x_1, \dots, x_m) \in \mathbb{R}_{\geq 0}^m$  un vecteur associant à la réaction  $r_i$  une vitesse  $x_i$ . Ce vecteur doit vérifier  $S_{\mathcal{J}}x = 0$ . Il est donc une combinaison des modes élémentaires (avec des coefficients réels positifs).

Géométriquement, l'ensemble de ces vecteurs forment un *cône*, que l'on note  $\text{Cône}(S_{\mathcal{J}}) = \{x \in \mathbb{R}_{\geq 0}^m \mid S_{\mathcal{J}}x = 0\}$ . Il y a alors un mode élémentaire pour chaque arête du cône (voir (Gagneur & Klamt, 2004)).

**Exemple 44.** Par exemple, reprenons les réactions R dans la Fig. 5.1, avec  $\mathcal{J} = \{X, Y\}$ . Tout vecteur du cône  $\text{Cône}(S_{\mathcal{J}})$  peut alors être obtenu par combinaison linéaire des deux modes élémentaires suivants, représentés à droite dans la figure :

$$x_1 = (1, 0, 0, 2) \quad x_2 = (0, 1, 1, 0)$$

On va alors montrer que les règles de simplification permettent de calculer les modes élémentaires d'un réseau. Intuitivement, l'élimination des espèces intermédiaires crée des modes, celle des flux dépendants élimine les modes qui sont la combinaison d'autres, et la factorisation permet d'obtenir des modes de taille minimale. On dit qu'un réseau  $U$  génère le cône  $C\hat{one}(S)$  si tout vecteur  $x \in C\hat{one}(S)$  est une combinaison linéaire positive de vecteurs de  $U$ .

**Lemme 9.** Soit  $U$  un réseau où les espèces  $\mathcal{J} \setminus \{X\}$  ont déjà été supprimées, et qui génère le cône  $C\hat{one}(S_{\mathcal{J} \setminus \{X\}})$ . Soit  $U'$  le réseau obtenu en appliquant la règle (INTERMÉDIAIRE) sur  $X$  dans  $U$ . Alors  $U'$  génère le cône  $C\hat{one}(S_{\mathcal{J}})$ .

*Proof.* On montre d'abord que les vecteurs de  $U'$  appartiennent bien au cône, c'est-à-dire que pour tout  $u \in U'$ , on a  $u \in C\hat{one}(S_{\mathcal{J}})$ . Puisque  $u \in U'$ , on a soit  $u \in U$ , soit il existe  $u_1, u_2 \in U$  avec  $u = u_1 \diamond_X u_2$ . Dans le premier cas, on a  $u \in C\hat{one}(S_{\mathcal{J} \setminus \{X\}})$  et  $stoic_{u,r}(X) = 0$ , donc  $u \in C\hat{one}(S_{\mathcal{J}})$ . Dans le second cas, on a  $u_1, u_2 \in C\hat{one}(S_{\mathcal{J} \setminus \{X\}})$ , donc  $u \in C\hat{one}(S_{\mathcal{J} \setminus \{X\}})$ . De plus,  $stoic_{u,r}(X) = 0$  par définition de  $\diamond_X$ , donc  $u \in C\hat{one}(S_{\mathcal{J}})$ .

On montre alors que tout vecteur  $x \in C\hat{one}(S_{\mathcal{J}})$  est une combinaison linéaire positive de vecteurs de  $U'$ . On sait que  $x \in C\hat{one}(S_{\mathcal{J} \setminus \{X\}})$ , et donc que  $x$  est la combinaison de vecteurs de  $U$ . On peut partitionner ces vecteurs selon leur stœchiométrie en  $X$ . Soit  $u_1, \dots, u_k \in U$  les vecteurs qui produisent du  $X$ , c'est-à-dire  $stoic_{u_i,r}(X) > 0$ . Soit  $u'_1, \dots, u'_{k'} \in U$  ceux qui consomment du  $X$ , c'est-à-dire  $stoic_{u'_i,r}(X) < 0$ , et  $u''_1, \dots, u''_{k''} \in U$  ceux où  $X$  est modificateur, c'est-à-dire  $stoic_{u''_i,r}(X) = 0$ . On a alors :

$$x = \sum_{1 \leq i \leq k} a_i u_i + \sum_{1 \leq i' \leq k'} a'_{i'} u'_{i'} + \sum_{1 \leq i'' \leq k''} a''_{i''} u''_{i''}.$$

De plus,  $stoic_{x,r}(X) = 0$ , donc  $\sum_{1 \leq i \leq k} a_i stoic_{u_i,r}(X) = - \sum_{1 \leq i' \leq k'} a'_{i'} stoic_{u'_{i'},r}(X)$ . On note

$D$  ce nombre dans la suite. On peut alors réécrire  $x$  de la façon suivante :

$$\begin{aligned} x &= \sum_{1 \leq i \leq k} a_i \frac{- \sum_{1 \leq i' \leq k'} a'_{i'} stoic_{u'_{i'},r}(X)}{D} u_i + \sum_{1 \leq i' \leq k'} a'_{i'} \frac{\sum_{1 \leq i \leq k} a_i stoic_{u_i,r}(X)}{D} u'_{i'} + \sum_{1 \leq i'' \leq k''} a''_{i''} u''_{i''} \\ &= \sum_{1 \leq i \leq k} \sum_{1 \leq i' \leq k'} \frac{a_i a'_{i'}}{D} (-stoic_{u'_{i'},r}(X)) u_i + \sum_{1 \leq i' \leq k'} \sum_{1 \leq i \leq k} \frac{a'_{i'} a_i}{D} stoic_{u_i,r}(X) u'_{i'} + \sum_{1 \leq i'' \leq k''} a''_{i''} u''_{i''} \\ &= \sum_{1 \leq i \leq k} \sum_{1 \leq i' \leq k'} \frac{a_i a'_{i'}}{D} (u_i \diamond_X u'_{i'}) + \sum_{1 \leq i'' \leq k''} a''_{i''} u''_{i''}. \end{aligned}$$

Comme  $X$  est un modificateur dans les vecteurs  $u''_{i''}$ , on a  $u''_{i''} \in U'$  pour tout  $i''$ . Et pour tout  $i, i'$ , on a  $u_i \diamond_X u'_{i'} \in U'$ . Donc  $x$  est bien la combinaison de vecteurs de  $U'$ , et donc  $U'$  génère le cône  $C\hat{one}(S_{\mathcal{J}})$ .  $\square$

On calcule alors les modes élémentaires d'un réseau en appliquant les règles de simplification. Soit  $\mathbf{r} = (r_1, \dots, r_m)$  une base de réactions, et  $S$  sa matrice de stœchiométrie. Soit  $U$  le réseau de flux initial, c'est-à-dire l'ensemble de vecteurs unitaires  $u_i$  tels que  $u_i \mathbf{r} = r_i$ . Soit  $\mathcal{J}$  l'ensemble des espèces intermédiaires.

**Théorème 5.** Soit  $U'$  le réseau obtenu en appliquant l'algorithme suivant, en partant de  $U$  et en terminant lorsque plus aucune règle de simplification n'est applicable.

Tant que  $\mathcal{J}$  est non vide :

1. appliquer la règle (**INTERMÉDIAIRE**) sur une espèce  $X \in \mathcal{J}$  encore présente dans le réseau,
2. supprimer tous les flux dépendants avec la règle (**DÉPENDANT**),
3. factoriser tous les flux avec la règle (**FACTORISATION**),
4.  $\mathcal{J} := \mathcal{J} \setminus \{X\}$ .

Alors  $U'$  est exactement l'ensemble des modes élémentaires de la matrice  $S_{\mathcal{J}}$ .

*Proof.* On procède par induction sur la taille de  $\mathcal{J}$ .

Si  $\mathcal{J} = \emptyset$ , alors  $S_{\mathcal{J}}$  est de taille 0, donc tout vecteur  $u \in \mathbb{N}^m$  vérifie  $S_{\mathcal{J}}u = 0$ . Et tout vecteur  $u$  est la combinaison des vecteurs unitaires de  $U$ , qui sont par définition minimaux. Donc  $U$  est bien l'ensemble des modes élémentaires.

Supposons que  $\mathcal{J}$  soit de taille  $k$ , et soit  $X \in \mathcal{J}$  une espèce intermédiaire. Par induction, on peut appliquer l'algorithme sur  $\mathcal{J} \setminus \{X\}$ , et on obtient un réseau  $U_1$  dont les flux sont les modes élémentaires de  $S_{\mathcal{J} \setminus \{X\}}$ . En particulier,  $U_1$  génère le cône  $\text{Cône}(S_{\mathcal{J} \setminus \{X\}})$ . Par le Lem. 9, en enlevant  $X$  avec la règle (**INTERMÉDIAIRE**), on obtient un réseau  $U_2$  qui génère le cône  $\text{Cône}(S_{\mathcal{J}})$ .

En particulier,  $U_2$  doit générer les arêtes du cône. Soit  $x \in \text{Cône}(S_{\mathcal{J}})$  un vecteur sur une des arêtes du cône. Pour qu'il soit généré par  $U_2$ , il faut alors qu'il existe un vecteur  $u \in U_2$  tel que  $x = au$ . Donc il existe dans  $U_2$  un vecteur pour chaque arête du cône.

Soit  $u \in U_2$  qui ne soit pas sur une arête. Alors  $u$  est la combinaison des vecteurs sur les arêtes, et peut être supprimé avec la règle (**DÉPENDANT**). On obtient un réseau  $U_3$  dont tous les vecteurs sont sur les arêtes.

En appliquant la règle (**FACTORISATION**) sur  $U_3$ , on obtient des vecteurs de tailles minimales. Donc  $U_3$  a un (unique) vecteur par arête du cône, de taille minimal, et l'ensemble de ces vecteurs génèrent tout le cône par combinaison linéaire. Donc ces vecteurs sont les modes élémentaires de la matrice  $S_{\mathcal{J}}$ .  $\square$

Cet algorithme est en fait similaire à la méthode de double description (voir par exemple ([Gagneur & Klamt, 2004](#))), qui permet de calculer les modes élémentaires d'une matrice.

**Exemple 45.** Reprenons les réactions de la Fig. 5.1. Après simplification, on obtient les réactions  $r_{23} = r_2 + r_3$  et  $r_{14^2} = r_1 + 2r_4$ , qui correspondent bien aux modes élémentaires représentés dans la Fig. 5.4.

### 5.1.5 Confluence

On montre maintenant que nos règles de simplification sont confluentes, et donc que quelque soit leur ordre d'application, on obtient toujours le même réseau. On a donc une méthode de simplification plus générale que la méthode de double description (Gagneur & Klamt, 2004), puisqu'il n'est pas nécessaire de supprimer les flux dépendants de façon systématique après chaque élimination d'espèces intermédiaires.

On prouve tout d'abord que l'on peut supprimer les espèces intermédiaires dans n'importe quel ordre. On prouve ensuite que les autres paires de règles de simplification commutent entre elles. Cela implique alors la confluence locale. Comme notre simplification termine trivialement, cela implique la confluence globale.

**Lemme 10.** Soit  $U$  un réseau, et  $X$  et  $Y$  deux espèces intermédiaires. Alors si on applique l'algorithme du Théoreme 5 en supprimant d'abord  $X$  puis  $Y$ , ou en supprimant d'abord  $Y$  puis  $X$ , on obtient le même réseau simplifié.

*Proof.* C'est la conséquence directe du Théoreme 5 et de l'unicité des modes élémentaires.  $\square$

**Lemme 11.** Les règles (DÉPENDANT) et (FACTORISATION) commutent ensemble, et commutent avec la règle (INTERMÉDIAIRE). C'est-à-dire que pour tout réseau  $U$ , pour toute règle  $D_1 \in \{Dep, Fact\}$  et toute règle  $D_2 \in \{Dep, Fact, Inter\}$ , si  $U \Rightarrow^{D_1} U_1$  et  $U \Rightarrow^{D_2} U_2$ , alors il existe un réseau  $U'$  tel que  $U_1 \Rightarrow^* U'$  et  $U_2 \Rightarrow^* U'$  (en utilisant les règles (DÉPENDANT), (FACTORISATION) et (INTERMÉDIAIRE)).

*Proof.* Soit  $U$  un réseau.

Supposons qu'on puisse simplifier  $U$  en  $U_1$  ou  $U_2$  en appliquant la règle (DÉPENDANT) respectivement sur un flux  $u = \sum_{1 \leq i \leq k} a_i u_i$  et  $u' = \sum_{1 \leq j \leq k'} a'_j u'_j$ , avec  $u_i, u'_j \in U$ . Si  $u = u'$ , alors  $U_1 = U_2$ . S'il existe  $i$  tel que  $u' = u_i$ , alors pour tout  $j$ , on a  $u \neq u'_j$ . On suppose (sans perte de généralité) que  $u' = u_1$ . On a alors  $u = \sum_{1 \leq j \leq k'} a'_j u'_j + \sum_{2 \leq i \leq k} a_i u_i$ . Donc, après avoir supprimé  $u'$ , on peut toujours supprimer  $u$ , et réciproquement, donc on obtient le même réseau simplifié. Dans tous les autres cas, les règles commutent trivialement.

Supposons que l'on puisse simplifier  $U$  en  $U_1$  en appliquant la règle (**DÉPENDANT**) sur un flux  $u = \sum_{1 \leq i \leq k} a_i u_i$ , avec  $u_i \in U$ , ou en  $U_2$  en appliquant (**FACTORISATION**) sur un flux  $a'u'$ . Si  $u = a'u'$  pour un entier  $a' > 1$ , alors pour tout  $i$  on a  $a_i u_i = a' a'_i u'_i$  pour des entiers  $a'_i$  et des flux  $u'_i$ . On peut alors appliquer (**FACTORISATION**) sur les  $u_i$  dans  $U_1$  et dans  $U_2$ , pour obtenir les flux  $u'_i$ , respectivement dans  $U'_1$  et  $U'_2$ . Dans  $U'_2$ , le flux  $u'$  est alors dépendant des flux  $u'_i$ , et peut donc être supprimé par application de (**DÉPENDANT**). On obtient alors le même réseau  $U'_1$ . Si  $u \neq a'u'$ , mais qu'il existe  $i$  tel que  $a'u' = u_i$ . On suppose, sans perte de généralité, que  $a'u' = u_1$ . Alors, dans  $U_1$ , on peut toujours factoriser  $a'u'$  en  $u'$  (avec (**FACTORISATION**)). Dans  $U_2$ , le flux  $u = \sum_{1 \leq i \leq k} a_i u_i = a_1 a'u' + \sum_{2 \leq i \leq k} a_i u_i$  est encore dépendant, et peut donc être supprimé avec la règle (**DÉPENDANT**). On obtient alors le même réseau simplifié. Dans tous les autres cas, les règles commutent trivialement.

Supposons maintenant qu'on simplifie  $U$  en  $U_1$  en appliquant (**DÉPENDANT**) sur  $u = \sum_{1 \leq i \leq k} a_i u_i$ , et en  $U_2$  en appliquant (**INTERMÉDIAIRE**) sur une espèce  $X$ . On peut partitionner les flux  $u_i$  en fonction de leur stœchiométrie en  $X$ . Sans perte de généralité, on suppose que les flux  $u_1$  à  $u_l$  produisent du  $X$ , ceux de  $u_{l+1}$  à  $u_{l'}$  en consomment, et ceux de  $u_{l'+1}$  à  $u_k$  ne modifient pas la quantité de  $X$ , pour des entiers  $1 \leq l \leq l' \leq k$ . Supposons dans un premier temps que  $\text{stoic}_{u_r}(X) = 0$ . Alors on a  $\sum_{1 \leq i \leq l} a_i \text{stoic}_{u_i r}(X) = \sum_{l+1 \leq j \leq l'} a_j (-\text{stoic}_{u_j r}(X))$ . On note  $D$  cet entier. On peut alors réécrire  $u$  de la façon suivante :

$$\begin{aligned}
u &= \sum_{1 \leq i \leq k} a_i u_i \\
&= 1/D \sum_{1 \leq i \leq l} D a_i u_i + 1/D \sum_{l+1 \leq j \leq l'} D a_j u_j + \sum_{l'+1 \leq i \leq k} a_i u_i \\
&= 1/D \sum_{1 \leq i \leq l} \sum_{l+1 \leq j \leq l'} a_j (-\text{stoic}_{u_j r}(X)) a_i u_i \\
&\quad + 1/D \sum_{l+1 \leq j \leq l'} \sum_{1 \leq i \leq l} a_i \text{stoic}_{u_i r}(X) a_j u_j + \sum_{l'+1 \leq i \leq k} a_i u_i \\
&= 1/D \sum_{1 \leq i \leq l} \sum_{l+1 \leq j \leq l'} a_i a_j (-\text{stoic}_{u_j r}(X)) u_i + \text{stoic}_{u_i r}(X) u_j + \sum_{l'+1 \leq i \leq k} a_i u_i \\
&= \sum_{1 \leq i \leq l} \sum_{l+1 \leq j \leq l'} a_i a_j / D (u_i \diamond_X u_j) + \sum_{l'+1 \leq i \leq k} a_i u_i.
\end{aligned}$$

Donc, dans  $U_2$ ,  $u$  est toujours dépendant, et peut être supprimé. Réciproquement, dans  $U_1$ , on peut toujours supprimer  $X$ , et on obtient dans les deux cas le même

réseau.

Supposons maintenant que  $stoic_{ur}(X) = a > 0$  (le cas  $a < 0$  est symétrique). Soit  $u' \in U$  un flux tel que  $stoic_{u'r}(X) = a' < 0$ . On va montrer que tout flux combinaison de  $u \diamond_X u'' \in U_2$  est dépendant et peut être supprimé. Ainsi, supprimer  $u$  puis  $X$  produit le même réseau que supprimer  $X$  puis tous les  $u \diamond_X u''$ . Comme  $stoic_{ur}(X) = a$ , on a  $\sum_{1 \leq i \leq l} a_i stoic_{u_i r}(X) = a + \sum_{l+1 \leq j \leq l'} a_j (-stoic_{u_j r}(X))$ . On note  $D$  cet entier. On a alors :

$$\begin{aligned}
u \diamond_X u' &= -a'u + au' \\
&= -a'/D \sum_{1 \leq i \leq l} Da_i u_i - a'/D \sum_{l+1 \leq j \leq l'} Da_j u_j - a' \sum_{l'+1 \leq i \leq k} a_i u_i + a/D(Du') \\
&= -a'/D \sum_{1 \leq i \leq l} aa_i u_i - a'/D \sum_{1 \leq i \leq l} \sum_{l+1 \leq j \leq l'} a_j (-stoic_{u_j r}(X)) a_i u_i \\
&\quad - a'/D \sum_{l+1 \leq j \leq l'} \sum_{1 \leq i \leq l} a_i stoic_{u_i r}(X) a_j u_j - a' \sum_{l'+1 \leq i \leq k} a_i u_i \\
&\quad + a/D \sum_{1 \leq i \leq l} a_i stoic_{u_i r}(X) u' \\
&= -a'/D \sum_{1 \leq i \leq l} \sum_{l+1 \leq j \leq l'} a_i u_j (-stoic_{u_j r}(X)) u_i + stoic_{u_i r}(X) u_j \\
&\quad + a/D \sum_{1 \leq i \leq l} a_i (-a' u_i + stoic_{u_i r}(X) u') - a' \sum_{l'+1 \leq i \leq k} a_i u_i \\
&= -a'/D \sum_{1 \leq i \leq l} \sum_{l+1 \leq j \leq l'} a_i a_j (u_i \diamond_X u_j) + a/D \sum_{1 \leq i \leq l} a_i (u_i \diamond_X u') - a' \sum_{l'+1 \leq i \leq k} a_i u_i.
\end{aligned}$$

Puisque  $a' < 0$ , les coefficients ci-dessus sont bien tous positifs. Alors  $u \diamond_X u'$  est bien dépendant dans  $U_2$ , et peut être supprimé, donc on obtient le même réseau simplifié.

Supposons que l'on puisse simplifier  $U$  en  $U_1$  ou  $U_2$  en appliquant la règle (**FACTORISATION**) respectivement sur un flux  $a_1 u_1$  et  $a_2 u_2$ . Si  $a_1 u_1 \neq a_2 u_2$ , les règles commutent trivialement : on peut toujours supprimer  $a_2 u_2$  dans  $U_1$  et symétriquement, et on obtient dans les deux cas le même réseau. Sinon, il existe  $a, a'_1, a'_2$  et un vecteur  $u$  tel que  $a_1 u_1 = a_2 u_2 = au$  et  $a = a_1 a'_1 = a_2 a'_2$ . Donc après avoir appliqué (**FACTORISATION**) sur  $a_1 u_1$ , on obtient le flux  $u_1 = a'_1 u$ , donc on peut encore appliquer la règle et obtenir le flux  $u$ . Par symétrie, on obtient le même flux si on commence par  $a_2 u_2$ .

Enfin, supposons qu'on puisse simplifier  $U$  en  $U_1$  en appliquant (**FACTORISATION**) sur un flux  $au$ , ou en  $U_2$  en appliquant (**INTERMÉDIAIRE**) sur une espèce  $X$ . Si  $stoic_{aur}(X) = 0$ , on peut toujours appliquer (**INTERMÉDIAIRE**) dans  $U_1$  et (**FACTORISATION**) dans  $U_2$ , et on obtient le même réseau simplifié. Sinon, supposons par exemple que  $stoic_{aur}(X) = a' > 0$  (l'autre cas est symétrique). Par définition, on



a  $a' = ab$  pour un entier  $b > 0$ . Dans  $U_2$ , on combine  $au$  avec tout flux  $u'$  tel que  $stoic_{u'}(X) = b' < 0$ . On a donc les flux  $au \diamond_X u' = b'au + abu' = a(b'u + bu')$ , qui peuvent être factorisés en  $b'u + bu'$ . Réciproquement, dans  $U_1$  on peut toujours appliquer (**INTERMÉDIAIRE**), et on obtient les flux  $u \diamond_X u' = b'u + bu'$ . Donc dans les deux cas on obtient le même réseau simplifié.  $\square$

La confluence est la conséquence directe des lemmes 10 et 11, et du lemme de Newmann (Lem. 1).

**Théorème 6.** La simplification est confluyente.

## 5.2 Confluence structurelle de la simplification avec cinétique

On considère maintenant les réseaux avec cinétique. On se restreint au cas de l'élimination d'espèces intermédiaires avec une cinétique linéaire (comme dans (INTERMÉDIAIRE - LINÉAIRE)). Afin de pouvoir supprimer l'ensemble des espèces intermédiaires, il est nécessaire que la linéarité soit préservée par la simplification, c'est-à-dire que si deux espèces intermédiaires  $X$  et  $Y$  sont linéaires, et que l'on supprime  $X$ , alors on veut que  $Y$  soit toujours linéaire dans le réseau simplifié. Pour cela, on montre qu'il est nécessaire d'ajouter des conditions supplémentaires sur les espèces intermédiaires, notamment sur leur stœchiométrie. Avec ces conditions, on montre qu'on peut toujours supprimer l'ensemble des espèces intermédiaires, et qu'on obtient une simplification structurellement confluente, c'est-à-dire que tous les réseaux complètement simplifiés ont tous les mêmes réactions si on néglige leurs expressions cinétiques.

### 5.2.1 Réseaux de flux cinétiques

Soit  $\mathbf{r} = (r_1, \dots, r_m)$  la base de réactions (sans cinétiques). Un *flux cinétique*  $v = u; e$  est une paire composée d'un flux  $u \in \mathbb{N}^m$  sur une base  $\mathbf{r}$  et d'une expression cinétique  $kin(v) = e$ . On note  $v\mathbf{r} = u\mathbf{r}; e$  la réaction cinétique associée. Un *réseau de flux cinétiques*  $W = V\&C$  est composé d'un ensemble  $V$  de flux cinétiques et d'une contrainte  $C$ . On note  $W\mathbf{r} = \{v\mathbf{r} \mid v \in V\}\&C$  le réseau de réactions correspondant. On étend naturellement la notion de similarité (Fig. 4.1) aux réseaux de flux, avec les deux axiomes de similarité présentés dans la Fig. 5.5.

**Exemple 46.** Considérons la base de réactions suivantes :

$$\mathbf{r} = \left( \begin{array}{ll} r_1 = A \rightarrow X, & r_4 = X \rightarrow Z, \\ r_2 = X \rightarrow Y, & r_5 = Y \rightarrow X + C, \\ r_3 = Y \rightarrow B, & r_6 = Z \rightarrow Y + D \end{array} \right).$$

On considère alors le réseau  $W$  représenté graphiquement dans la Fig. 5.6 (en haut à gauche). Ce réseau possède trois espèces intermédiaires  $\{X, Y, Z\}$ , que l'on suppose à l'équilibre. Notons que pour simplifier les réseaux, on ne fait pas apparaître les flèches en pointillé indiquant les modificateurs, puisque l'information est déjà présente dans les cinétiques. On procédera de la même façon pour toutes les figures de ce chapitre.

La réaction  $r_1$  transforme une molécule de  $A$  en une molécule de l'espèce intermédiaire  $X$ , qui peut alors se transformer soit en  $Y$  ( $r_2$ ), soit en  $Z$  ( $r_4$ ). La molécule de  $Y$  peut soit se transformer en  $B$  ( $r_3$ ), soit se retransformer en  $X$ , en

Réseaux de flux cinétiques

$$\frac{C \vdash e \equiv e'}{V \cup \{u; e\} \& C \equiv V \cup \{u; e'\} \& C} \quad \frac{C \equiv C'}{V \& C \equiv V \& C'}$$

Fig 5.5: Axiomes de similarité pour les réseaux de flux cinétiques.

produisant dans le même temps un molécule de  $C$  ( $r_5$ ). Une molécule de  $Z$  peut se transformer en une molécule de  $Y$ , en produisant également une molécule de  $D$  ( $r_6$ ). Toutes les cinétiques sont des lois d'action de masse, et sont linéaires dans les concentrations des espèces intermédiaires.

### 5.2.2 Espèces stables-linéaires

On cherche à supprimer un ensemble d'espèces intermédiaires, noté  $\mathcal{J} \subseteq \mathcal{I}$ . Pour cela, on se restreint ici au cas d'espèces intermédiaires stables et avec une des expressions cinétiques linéaires, comme défini ci-dessous.

On note  $Prod_{\mathcal{J}}(\mathbf{vr}) = \{X \in \mathcal{J} \mid stoic_{\mathbf{vr}}(X) > 0\}$  et  $Reac_{\mathcal{J}}(\mathbf{vr}) = \{X \in \mathcal{J} \mid stoic_{\mathbf{vr}}(X) < 0\}$  les ensembles d'espèces intermédiaires produites et consommées par une réaction  $\mathbf{vr}$ .

**Définition 21.** Une espèce  $X$  est *stable-linéaire* dans un réseau de flux  $V \& C$  si elle vérifie les quatre conditions suivantes :

- **Stabilité** :  $C \equiv C' \wedge cst(x_X)$ .
- **Consommation linéaire** : l'expression cinétique de tout flux qui consomme  $X$  est linéaire en  $x_X$  : pour tout  $v \in V$  tel que  $X \in Reac_{\mathcal{J}}(\mathbf{vr})$ ,  $kin(v) \equiv x_X e$ , avec une expression  $e$  telle que  $x_X \notin Vars(e)$ .
- **Production indépendante** : L'expression cinétique de tout flux qui produit  $X$  ne dépend pas de  $x_X$  : pour tout  $v \in V$  tel que  $X \in Prod_{\mathcal{J}}(\mathbf{vr})$ ,  $x_X \notin Vars(kin(v))$ .
- **Consommation non nulle** : Elle est toujours consommée avec une cinétique non nulle, c'est-à-dire, en notant  $V_{reac} = \{v \in V \mid X \in Reac_{\mathcal{J}}(\mathbf{vr})\}$  :  $C \equiv C'' \wedge \sum_{v \in V_{reac}} kin(v) \neq 0$ .

Supposons qu'une espèce  $X$  soit stable-linéaire dans  $W = V \& C$ . La première condition implique  $\dot{x}_X = 0$ . Les deux conditions suivantes impliquent  $\dot{x}_X = T - x_X T'$ , avec  $x_X \notin Vars(T), Vars(T')$ , où  $T$  (respectivement  $x_X T'$ ) est la somme des

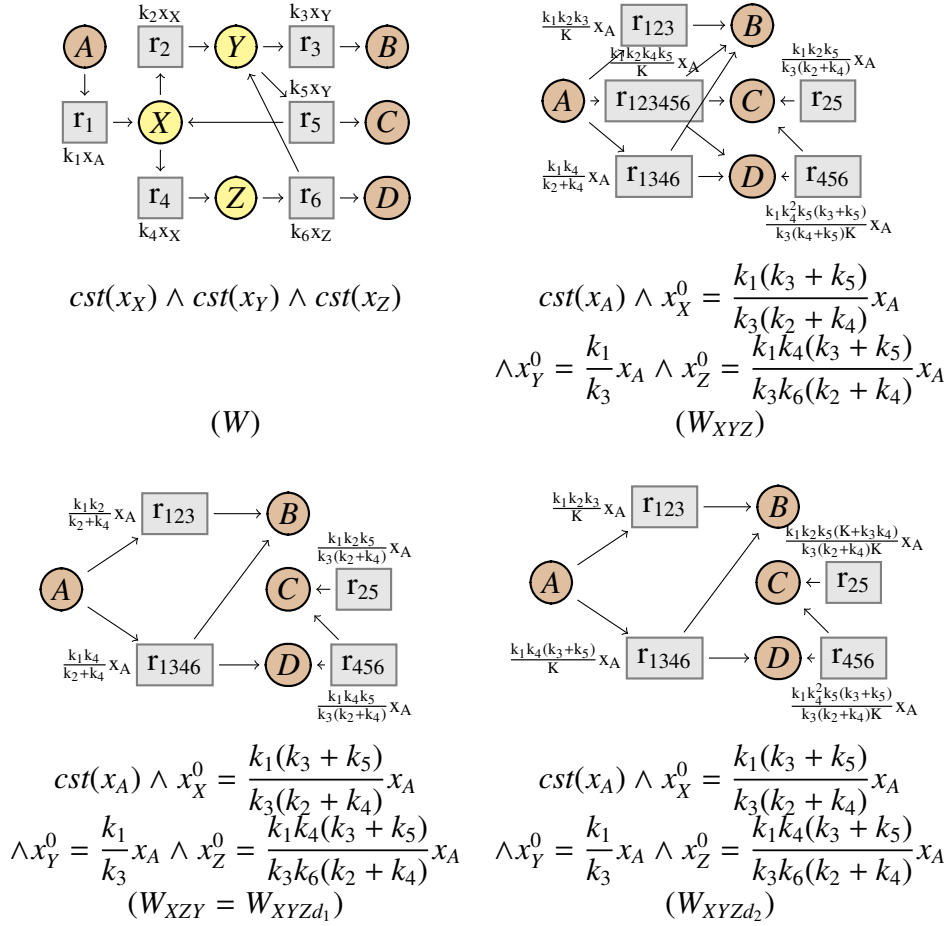


Fig 5.6: Réseau W (en haut et à gauche) et ses simplifications : en haut à droite le réseau W<sub>XYZ</sub> obtenu en supprimant dans l'ordre les espèces intermédiaires X, Y et Z; en bas à gauche, le réseau W<sub>XZY</sub> obtenu en supprimant X puis Z puis Y; en supprimant le flux dépendant  $u_{123456}$  du réseau W<sub>XYZ</sub>, on peut obtenir soit le réseau W<sub>XYZd1</sub> = W<sub>XZY</sub>, soit le réseau W<sub>XYZd2</sub> en bas à droite. Ces deux réseaux ont la même structure, mais pas les mêmes expressions cinétiques. Le nouveau paramètre K dans les expressions cinétiques est défini par  $K = k_2 k_3 + k_3 k_4 + k_4 k_5$ .

cinétiques des flux qui produisent (respectivement consomment)  $X$ . La dernière condition implique  $T' \neq 0$ . On peut alors en déduire  $x_X = T/T'$ .

Remarquons qu'on peut par contre avoir  $T = 0$ , et donc  $x_X = 0$ . Dans ce cas, la règle de simplification d'une espèce intermédiaire consiste à supprimer les réactions qui consomment  $X$ , puisque sa concentration est toujours nulle.

**Exemple 47.** Dans le réseau  $W$  illustré en haut à gauche de la Fig. 5.6, toutes les espèces intermédiaires  $\mathcal{J} = \{X, Y, Z\}$  sont stables-linéaires.

### 5.2.3 Simplification

On présente alors les quatre règles de simplification pour les réseaux de flux cinétiques dans la Fig. 5.7.

La règle (**INTERMÉDIAIRE**) supprime une espèce intermédiaire stable-linéaire  $X$ , en combinant les flux deux à deux avec l'opérateur suivant, où  $v = u$ ;  $e$  est un flux produisant  $X$  et  $v' = u'$ ;  $x_X e'$  un flux consommant  $X$  :

$$v \diamond_{T'} v' = (-\text{stoic}_{v'r}(X))u + (\text{stoic}_{v'r}(X))u'; ee'/T'.$$

La règle (**MODIFICATEUR CONSTANT**) supprime une espèce intermédiaire qui n'est plus présente ni en réactant ni produit de en réactions, et qui est donc juste un modificateur.

La règle (**DÉPENDANT**) supprime un flux  $v = u$ ;  $e$  qui dépend d'un ensemble de flux  $u_i$ ;  $e_i$ , c'est-à-dire tel que  $u = \sum_{1 \leq i \leq k} a_i u_i$ . Dans ce cas, la cinétique  $e_i$  de ces flux est également remplacée par  $e_i + a_i e$ .

La règle (**SIMILARITÉ**) fonctionne de la façon habituelle.

La règle de factorisation n'est plus nécessaire ici, grâce aux conditions de linéarité et de stœchiométrie. On définit également la règle de similarité, comme pour les réseaux de réactions cinétiques. Comme pour le cas sans cinétique, les règles sont définies de façon globale, et il n'y a donc pas besoin d'une règle pour le contexte.

Il y a alors une relation de simulation entre la simplification de réseaux de flux cinétiques et la simplification de réseaux de réactions cinétiques.

**Proposition 14.** Soit  $\mathbf{r}$  une base de réactions, et  $W, W'$  deux réseaux de flux tels que  $W \Rightarrow W'$ . Alors  $W\mathbf{r} \Rightarrow W'\mathbf{r}$ .

*Proof.* La règle (**INTERMÉDIAIRE**) peut être imitée par (**INTERMÉDIAIRE - LINÉAIRE**) et la règle (**DÉPENDANT**) par (**COMBINAISON**). Les règles (**MODIFICATEUR CONSTANT**) et (**SIMILARITÉ**) peuvent être imitées par les règles équivalentes dans les réseaux de réactions cinétiques.  $\square$

(INTERMÉDIAIRE)

$$\begin{array}{c}
X \in \mathcal{J} \quad X \text{ est stable-linéaire dans } V\&C \\
\frac{T = \sum_{\{u; e \in V \mid X \in \text{Prod}_{\mathcal{J}}(u\mathbf{r})\}} e \quad T' = \sum_{\{u'; x_X e' \in V \mid X \in \text{Reac}_{\mathcal{J}}(u'\mathbf{r})\}} e' \quad F = T/T'}{\frac{\{v \diamond_{T'} v' \mid v, v' \in V, \text{stoic}_{v\mathbf{r}}(X) > 0, \text{stoic}_{v'\mathbf{r}}(X) < 0\}}{V\&C \Rightarrow^{\text{Inter}} \cup \{v[x_X := F] \mid v \in V, \text{stoic}_{v\mathbf{r}}(X) = 0\}} \\
& \&C[x_X := F] \wedge x_X^0 = F
\end{array}$$

(MODIFICATEUR CONSTANT)

$$\frac{X \in \mathcal{J} \quad \forall v \in V. \text{stoic}_{v\mathbf{r}}(X) = 0}{V\&C \Rightarrow^{\text{Mod}} V[x_X := x_X^0] \&C[x_X := x_X^0]}$$

(DÉPENDANT)

$$\frac{a_1, \dots, a_k \in \mathbb{N}_{>0} \quad u = \sum_{1 \leq i \leq k} a_i u_i}{\{u_i; e_i \mid 1 \leq i \leq k\} \cup \{u; e\} \cup V \Rightarrow^{\text{Dep}} \{u_i; e_i + a_i e \mid 1 \leq i \leq k\} \cup V}$$

(SIMILARITÉ)

$$\frac{W_1 \equiv W'_1 \quad W'_1 \Rightarrow W'_2 \quad W'_2 \equiv W_2}{W_1 \Rightarrow^{\text{Sim}} W_2}$$

Fig 5.7: Règles de simplification de réseaux de flux cinétiques.

Deux réseaux de flux  $W, W'$  sont congruents si les réseaux de réactions  $W\mathbf{r}, W'\mathbf{r}$  correspondants sont congruents, c'est-à-dire que dans tout contexte, leurs systèmes d'équations ont les mêmes solutions observables.

**Définition 22.** Soit  $\mathbf{r}$  une base de réactions,  $\Omega$  une fonction d'observation,  $\mathcal{I}$  des espèces internes. Deux réseaux de flux  $W$  et  $W'$  sont  $(\Omega, \mathcal{I})$ -congruents, noté  $W \sim_{\Omega, \mathcal{I}} W'$ , si  $W\mathbf{r} \sim_{\Omega, \mathcal{I}} W'\mathbf{r}$ .

Comme l'observation et les espèces internes sont en général fixées, on notera la congruence  $\sim$ .

**Proposition 15.** La simplification de réseaux de flux est correcte pour la congruence :  $W \Rightarrow W'$  implique  $W \sim W'$ .

*Proof.*  $W \Rightarrow W'$  implique  $W\mathbf{r} \Rightarrow W'\mathbf{r}$  (Prop. 14), qui implique  $W\mathbf{r} \sim W'\mathbf{r}$  (Théoreme 4), et donc  $W \sim W'$  (par définition).  $\square$

### 5.2.4 Réseaux linéaires

Afin de pouvoir simplifier l'ensemble des espèces intermédiaires  $\mathcal{J}$ , il faut que pour tout  $X \in \mathcal{J}$ , soit  $X$  soit stable-linéaire dans  $W$ , et donc simplifiable avec la règle (**INTERMÉDIAIRE**), soit que  $X$  soit un modificateur dans  $W$ , et donc simplifiable avec (**MODIFICATEUR CONSTANT**). Mais il faut également que lorsque l'on simplifie une espèce intermédiaire  $X$ , les autres espèces vérifient toujours ces conditions dans le réseau simplifié. Or ce n'est pas le cas en général (voir l'Ex. 49 ci-dessous), et il est donc nécessaire de se limiter à une sous-classe de réseaux de flux, notée *LinNets*.

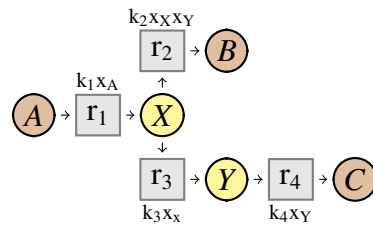
**Définition 23.** Soit  $\mathcal{J} \subseteq \mathcal{I}$  un ensemble d'espèces intermédiaires (internes). Un réseau  $W = V\&C$  appartient à l'ensemble *LinNets* si, pour toute espèce  $X \in \mathcal{J}$ , on a :

1. Soit  $X$  est stable-linéaire dans  $W$ , soit  $X$  est uniquement un modificateur, c'est-à-dire pour tout  $v; e \in V$ ,  $stoic_{v\mathbf{r}}(X) = 0$ .
2. La cinétique de toute réaction qui consomme  $X$  ne dépend pas d'autres espèces intermédiaires : pour tout  $v \in V$  tel que  $X \in Reac_{\mathcal{J}}(v\mathbf{r})$ , pour tout  $Y \in \mathcal{J}$  telle que  $Y \neq X$ , on a  $Y \notin Espèces(kin(v))$ .
3. La cinétique de toute réaction qui produit  $X$  et ne consomme pas d'espèces intermédiaires ne dépend pas des espèces intermédiaires : pour tout  $v \in V$  tel que  $X \in Prod_{\mathcal{J}}(v\mathbf{r})$  et  $Reac_{\mathcal{J}}(v\mathbf{r}) = \emptyset$ , on a  $\mathcal{J} \cap Espèces(kin(v)) = \emptyset$ .
4. La stœchiométrie totale des espèces intermédiaires en réactant (resp. produit) d'une réaction est toujours inférieure à 1 : ainsi, pour tout  $v \in V$ ,  $|Reac_{\mathcal{J}}(v\mathbf{r})| \leq 1$  et  $|Prod_{\mathcal{J}}(v\mathbf{r})| \leq 1$ . Donc d'une part, il ne peut jamais qu'une seule espèce intermédiaire dans les réactants (ou les produits) d'une réaction. Et d'autre part, il ne peut pas y avoir plus d'une molécule de cette espèce intermédiaire.

**Exemple 48.** Dans l'Ex. 46 de la Fig. 5.6, le réseau  $W$  appartient bien à *LinNets*, avec  $\mathcal{J} = \{X, Y, Z\}$ .

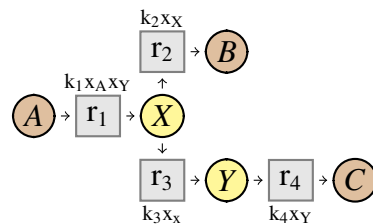
L'exemple suivant illustre les motivations pour ces conditions.

**Exemple 49.** Considérons tout d'abord un réseau dans lequel une réaction consomme  $X$  avec une cinétique dépendant d'une autre espèce intermédiaire (ici  $Y$  dans la réaction  $r_2$ ), c'est-à-dire tel que la condition 2 n'est pas respectée :



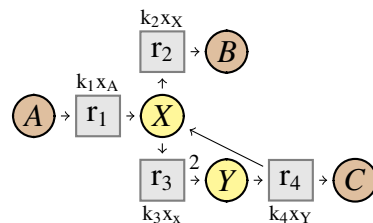
Si on supprime l'espèce  $Y$  en première, on combine les réactions  $r_3$  et  $r_4$ , et on calcule  $x_Y = \frac{k_3}{k_4}x_X$ . On remplace alors  $x_Y$  par cette valeur dans  $r_2$ , qui a alors une cinétique non linéaire en  $x_X$  :  $\frac{k_2k_3}{k_4}x_X^2$ .

De façon similaire, supposons qu'une réaction produit  $X$ , ne consomme pas  $Y$ , mais a une cinétique qui dépend de  $x_Y$  (la réaction  $r_1$  ci-dessous), donc un réseau dans lequel la condition 3 n'est pas respectée :



Si on enlève  $Y$ , l'expression de  $r_1$  devient  $\frac{k_1k_3}{k_4}x_Ax_X$ . On a donc une réaction produisant  $X$  avec une cinétique qui dépend de  $x_X$ . Donc l'équation différentielle  $\dot{x}_X = x_X(\frac{k_1k_3}{k_4}x_A - (k_2 + k_3))$  n'a pas la bonne forme, et on ne peut pas calculer  $x_X$ .

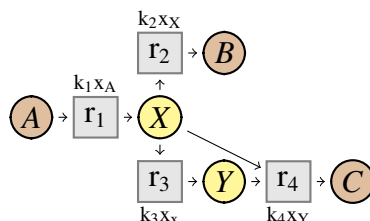
Le même problème peut apparaître si le réseau ne vérifie pas la condition 4, par exemple si une réaction produit une espèce intermédiaire avec une stœchiométrie de 2 :



Si on supprime  $Y$ , la réaction  $r_3 \diamond_{k_4} r_4$  produit une molécule de  $X$  avec une cinétique  $k_3x_X$ .



Enfin, si deux espèces intermédiaire interagissent ensemble en réactants ou en produits d'une réaction (condition 4), alors la stœchiométrie d'un des intermédiaires peut devenir supérieur strictement à 1 :



Si on enlève  $Y$ , la combinaison des réactions  $r_3$  et  $r_4$  est une réaction qui consomme deux molécules de  $X$ .

### 5.2.5 Caractérisation des dépendances

On veut alors montrer que l'ensemble  $LinNets$  est bien stable par simplification, c'est-à-dire que si  $W \in LinNets$  et  $W \Rightarrow^* W'$ , alors  $W' \in LinNets$ . Pour cela, on montre tout d'abord dans cette sous-section la Prop. 16 suivante, qui décrit certaines propriétés des flux dépendants.

**Proposition 16.** Soit  $v_d \in W$  un flux dépendant de  $v_1, \dots, v_k$ , avec les coefficients  $a_1, \dots, a_k$ . Si  $Reac_{\mathcal{J}}(v_d \mathbf{r}) = Prod_{\mathcal{J}}(v_d \mathbf{r}) = \emptyset$ , alors pour tout  $1 \leq j \leq k$ ,  $Reac_{\mathcal{J}}(v_j \mathbf{r}) = Prod_{\mathcal{J}}(v_j \mathbf{r}) = \emptyset$ . Sinon, il existe un unique index  $1 \leq i \leq k$  tel que  $Reac_{\mathcal{J}}(v_i \mathbf{r}) = Reac_{\mathcal{J}}(v_d \mathbf{r})$  et  $Prod_{\mathcal{J}}(v_i \mathbf{r}) = Prod_{\mathcal{J}}(v_d \mathbf{r})$ . De plus, dans ce cas, pour tout  $j \neq i$ ,  $Reac_{\mathcal{J}}(v_j \mathbf{r}) = Prod_{\mathcal{J}}(v_j \mathbf{r}) = \emptyset$ .

Afin de prouver cette proposition, il est nécessaire d'introduire une notion de chemin. Notons que toutes les définitions introduites ici ne sont utiles que pour prouver la Prop. 16, et ne seront pas utilisées ailleurs.

**Définition 24.** Soit  $\mathbf{r}$  la base initiale de réactions, et  $\mathcal{J}$  l'ensemble des espèces intermédiaires. On définit les notions suivantes :

- un *chemin*  $\tilde{u} = u_1 \dots u_k$  est une séquence non vide de vecteurs unitaires  $u_i$ , telle que pour tout  $1 \leq i < k$ , on a  $Prod_{\mathcal{J}}(u_i \mathbf{r}) = Reac_{\mathcal{J}}(u_{i+1} \mathbf{r}) \neq \emptyset$ ,
- le *vecteur* d'un chemin est défini par  $\sum \tilde{u} = \sum_{1 \leq i \leq k} u_i$ ,
- un chemin est *circulaire* si  $Prod_{\mathcal{J}}(u_k \mathbf{r}) = Reac_{\mathcal{J}}(u_1 \mathbf{r}) \neq \emptyset$ , et *non-circulaire* sinon,

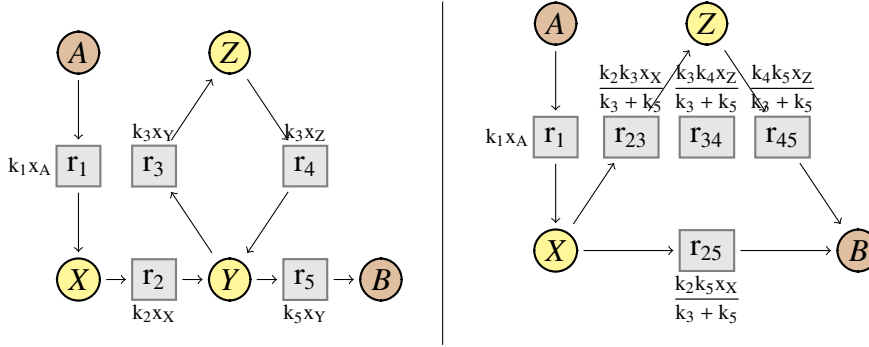


Fig 5.8: Le réseau initial  $W_0$  (gauche) et le réseau simplifié  $W$  (droite) après avoir enlevé  $Y$ .

- pour un chemin circulaire, on note  $\tilde{u}(X) = |\{1 \leq i \leq k \mid \text{Prod}_{\mathcal{J}}(u_i \mathbf{r}) = \{X\}\}|$  le multi-ensemble d'espèces intermédiaires traversées par  $\tilde{u}$ ,
- pour un chemin non-circulaire  $\tilde{u}$ , on note  $\text{Reac}_{\mathcal{J}}(\tilde{u} \mathbf{r}) = \text{Reac}_{\mathcal{J}}(u_1 \mathbf{r})$  son début et  $\text{Prod}_{\mathcal{J}}(\tilde{u} \mathbf{r}) = \text{Prod}_{\mathcal{J}}(u_k \mathbf{r})$  sa fin. On définit également le multi-ensemble  $\tilde{u}(X) = |\{1 \leq i < k \mid \text{Prod}_{\mathcal{J}}(u_i \mathbf{r}) = \text{Reac}_{\mathcal{J}}(u_{i+1} \mathbf{r}) = \{X\}\}|$ . Notons qu'on ne compte ni  $\text{Reac}_{\mathcal{J}}(\tilde{u} \mathbf{r})$  ni  $\text{Prod}_{\mathcal{J}}(\tilde{u} \mathbf{r})$  dans ce multi-ensemble.

**Exemple 50.** Considérons le réseau  $W_0$  à gauche de la Fig. 5.8, avec  $\mathcal{J} = \{X, Y, Z\}$ . On dénote par  $u_i$  le vecteur unitaire pour la réaction  $\mathbf{r}_i$ .

Le chemin  $\tilde{u} = u_2 u_5$  est non-circulaire. Son début est  $\text{Reac}_{\mathcal{J}}(\tilde{u} \mathbf{r}) = \{X\}$  et sa fin  $\text{Prod}_{\mathcal{J}}(\tilde{u} \mathbf{r}) = \emptyset$ . Son vecteur est  $\sum \tilde{u} = (0, 1, 0, 0, 1)$ . Le multi-ensemble correspondant est défini par  $\tilde{u}(X) = \tilde{u}(Z) = 0$  (puisque seules les étapes intermédiaires sont comptées dans ce multi-ensemble pour les chemins non-circulaires) et  $\tilde{u}(Y) = 1$ .

Le chemin  $\tilde{u}' = u_1 u_2 u_3 u_4$  est aussi non-circulaire, avec  $\text{Reac}_{\mathcal{J}}(\tilde{u}' \mathbf{r}) = \text{Prod}_{\mathcal{J}}(\tilde{u}' \mathbf{r}) = \emptyset$  et  $\sum \tilde{u}' = (1, 1, 1, 1, 1)$ . Son multi-ensemble est défini par  $\tilde{u}'(X) = \tilde{u}'(Z) = 1$  et  $\tilde{u}'(Y) = 2$ .

Le chemin  $\tilde{u}'' = u_3 u_4$  est circulaire, avec pour vecteur  $\sum \tilde{u}'' = (0, 0, 1, 1, 0)$ . Son multi-ensemble est défini par  $\tilde{u}''(X) = 0$  et  $\tilde{u}''(Y) = \tilde{u}''(Z) = 1$ .

On va montrer plus loin que si un chemin  $\tilde{u}$  vérifie certaines conditions par rapport à un réseau  $W$ , alors il existe un flux  $v = u; e \in W$  tel que  $\sum \tilde{u} = u$ , et réciproquement. Ces chemins particuliers, appelés *chemins de flux*, sont définis ci-dessous. Un chemin de flux peut être soit circulaire, soit non-circulaire.

**Définition 25.** Soit  $W_0$  le réseau initial et  $W$  un réseau simplifié, c'est-à-dire  $W_0 \Rightarrow^* W$ . On définit alors :

- un *chemin de flux non-circulaire*  $\tilde{u}$  dans  $W$  est un chemin non-circulaire tel que :
  - pour toute espèce intermédiaire  $X$  vérifiant que  $\tilde{u}(X) > 0$ , on a  $X \in \text{Espèces}(W_0) \setminus \text{Espèces}(W)$  (c'est-à-dire que  $X$  a été enlevée par l'une des étapes de simplification dans  $W_0 \Rightarrow^* W$ ),
  - $\text{Reac}_{\mathcal{J}}(\tilde{u}\mathbf{r})$  est soit l'ensemble vide  $\emptyset$ , soit une espèce  $X \in \text{Espèces}(W)$  qui n'a pas encore été supprimée dans  $W$ , et similairement pour  $\text{Prod}_{\mathcal{J}}(\tilde{u}\mathbf{r})$ ,
- un *chemin de flux circulaire*  $\tilde{u}$  dans  $W$  est un chemin circulaire s'il y a au maximum une espèce intermédiaire  $X \in W$  (c'est-à-dire qui n'a pas encore été supprimée) telle que  $\tilde{u}(X) > 0$ ,
- un chemin de flux  $\tilde{u}$  correspond à un flux  $v = u; e$  si  $\sum \tilde{u} = u$ .

**Exemple 51.** Considérons le réseau simplifié  $W$  après avoir enlevé l'espèce intermédiaire  $Y$  du réseau  $W_0$  (Fig. 5.8).

Le chemin  $\tilde{u} = u_2u_5$  est un chemin de flux non-circulaire. Il correspond au flux  $v_{25}$ . Le chemin  $\tilde{u}' = u_1u_2u_3u_4u_5$  n'est pas un chemin de flux, puisque  $\tilde{u}'(X) > 0$  et  $X$  n'a pas encore été supprimée. Le chemin  $\tilde{u}'' = u_3u_4$  est un chemin de flux circulaire, puisqu'il existe une unique espèce intermédiaire,  $Z$ , telle que  $Z \in W$  et  $\tilde{u}''(Z) > 0$ . Il correspond au flux  $v_{34}$ .

**Lemme 12.** Soit  $\tilde{u}$  un chemin de flux circulaire correspondant au flux  $v$ . Alors  $\text{Reac}_{\mathcal{J}}(v\mathbf{r}) = \text{Prod}_{\mathcal{J}}(v\mathbf{r}) = \emptyset$ .

*Proof.* Soit  $\tilde{u} = u_1 \dots u_k$  un chemin de flux circulaire correspondant au flux  $v$ . Pour tout  $1 \leq i < k$  on a  $\text{Prod}_{\mathcal{J}}(u_i\mathbf{r}) = \text{Reac}_{\mathcal{J}}(u_{i+1}\mathbf{r})$ . De plus,  $\text{Prod}_{\mathcal{J}}(u_k\mathbf{r}) = \text{Reac}_{\mathcal{J}}(u_1\mathbf{r})$ . On a donc, pour toute espèce intermédiaire  $X$ , on a :

$$\text{stoic}_{v\mathbf{r}}(X) = \sum_{1 \leq i \leq k} \text{stoic}_{u_i\mathbf{r}}(X) = 0.$$

Donc  $\text{Reac}_{\mathcal{J}}(v\mathbf{r}) = \text{Prod}_{\mathcal{J}}(v\mathbf{r}) = \emptyset$ . □

On prouve alors le lemme clé de cette section.

**Lemme 13.** Soit  $W_0$  le réseau initial et  $W$  un réseau simplifié tel que  $W_0 \Rightarrow^* W$ . Alors les propriétés suivantes sont vérifiées :

1. pour tout flux  $v \in W$ , il existe un chemin de flux  $\tilde{u}$  pour  $W$ , correspondant à  $v$ . De plus, si  $v$  n'est pas dépendant dans  $W$ , alors pour toute espèce intermédiaire  $X$ , on a  $\tilde{u}(X) \leq 1$ ,

2. pour tout chemin de flux  $\tilde{u}$  pour  $W$  tel que pour toute espèce intermédiaire  $X$ ,  $\tilde{u}(X) \leq 1$ , il existe un flux  $v \in W$  correspondant,
3. soit  $v_d \in W$  un flux qui dépend de  $v_1, \dots, v_k \in W$  avec les coefficients  $a_1, \dots, a_k$ . Si  $Reac_{\mathcal{J}}(v_d \mathbf{r}) = Prod_{\mathcal{J}}(v_d \mathbf{r}) = \emptyset$ , alors pour tout  $1 \leq i \leq k$ ,  $Reac_{\mathcal{J}}(v_j \mathbf{r}) = Prod_{\mathcal{J}}(v_j \mathbf{r}) = \emptyset$ . Sinon :
  - il existe un unique index  $1 \leq i \leq k$  tel que  $a_i = 1$ ,  $Reac_{\mathcal{J}}(v_i \mathbf{r}) = Reac_{\mathcal{J}}(v_d \mathbf{r})$  et  $Prod_{\mathcal{J}}(v_i \mathbf{r}) = Prod_{\mathcal{J}}(v_d \mathbf{r})$ ,
  - pour tout  $j \neq i$ ,  $Reac_{\mathcal{J}}(v_j \mathbf{r}) = Prod_{\mathcal{J}}(v_j \mathbf{r}) = \emptyset$ , et tout chemin de flux  $\tilde{u}_j$  correspondant à  $v_j$  est circulaire.

*Proof.* On procède par induction sur les étapes de simplification. On considère tout d'abord le cas  $W = W_0$ .

- (1) pour tout flux  $v = u; e$  dans le réseau initial  $W_0$ , le vecteur  $u$  est forcément un vecteur unitaire  $u_i$  pour la base  $\mathbf{r}$ . Donc on peut directement lui associer le chemin de flux correspondant  $\tilde{u} = u_i$ . Puisque c'est un vecteur unitaire, le flux  $v$  est également forcément non dépendant. Un chemin de flux de taille 1 est obligatoirement non-circulaire, et vérifie  $\tilde{u}(X) = 0$  pour toute espèce intermédiaire  $X$ , donc on a bien  $\tilde{u}(X) \leq 1$ .
- (2) un chemin de flux  $\tilde{u}$  pour  $W_0$  est forcément de taille 1. En effet, dans le cas contraire, si  $\tilde{u} = u_1 \dots u_k$  avec  $k > 1$  est non-circulaire, alors il existe une espèce intermédiaire  $X \in Prod_{\mathcal{J}}(u_1 \mathbf{r})$  avec  $X \in Espèces(W_0) \setminus Espèces(W_0) = \emptyset$ . Si  $\tilde{u}$  est circulaire, alors il existe  $X \in Prod_{\mathcal{J}}(u_1 \mathbf{r})$ ,  $Y \in Prod_{\mathcal{J}}(u_2 \mathbf{r})$  avec  $\tilde{u}(X) > 0$  et  $\tilde{u}(Y) > 0$ , ce qui contredit la définition de chemin de flux circulaire. Donc  $\tilde{u}$  est de taille 1, et donc il existe  $i$  tel que  $\tilde{u} = u_i$ , et un flux  $u_i; e \in W_0$  correspondant,
- (3) comme précisé précédemment, un flux  $v \in W_0$  ne peut pas être dépendant.

Considérons maintenant l'étape d'induction. On suppose que  $W_0 \Rightarrow^k W'$  (pour  $k > 0$ ), que le lemme est vérifié pour  $W'$ , et que  $W' \Rightarrow W$ . Si  $W' \Rightarrow_{Mod} W$  ou  $W' \Rightarrow_{Sim} W$ , seules les expressions cinétiques sont modifiées entre  $W'$  et  $W$ , et donc le lemme est toujours vérifié dans  $W$ .

On considère maintenant le cas  $W' \Rightarrow_{Dep} W$ . On prouve le lemme point par point.

- (1) Soit  $v = u; e \in W$  un flux. La règle (**DÉPENDANT**) supprime un flux et modifie certaines expressions cinétiques. Donc il existe un flux  $v' = u; e' \in W'$ .

Par hypothèse d'induction, il existe un chemin de flux  $\tilde{u}$  pour  $W'$ , correspondant à  $v'$ , c'est-à-dire tel que  $\sum \tilde{u} = u$ . Comme de plus  $\text{Espèces}(W') = \text{Espèces}(W)$ ,  $\tilde{u}$  est encore un chemin de flux pour  $W$ , et correspond à  $v$ . De plus, supposons que  $v'$  soit dépendant dans  $W'$ . Soit  $v_d$  le flux dépendant éliminé par la règle de simplification. Alors, en particulier,  $v'$  est dépendant d'un ensemble de flux ne contenant pas  $v_d$ . En effet, si  $v'$  dépend de  $v_d$  et de  $v_1, \dots, v_k$ , et que  $v_d$  dépend de  $v'_1, \dots, v'_l$ , alors  $v'$  est également dépendant de  $v_1, \dots, v_k, v'_1, \dots, v'_l$ . Donc, en particulier, si  $v'$  est dépendant dans  $W'$ , alors  $v$  est dépendant dans  $W$ . Et si  $v$  n'est pas dépendant,  $v'$  non plus, et par induction pour toute espèce intermédiaire  $X$ , on a  $\tilde{u}(X) \leq 1$ .

- (2) Soit  $\tilde{u}$  un chemin de flux pour  $W$  tel que pour toute espèce intermédiaire  $X$ , on ait  $\tilde{u}(X) \leq 1$ . Comme  $\text{Espèces}(W) = \text{Espèces}(W')$ ,  $\tilde{u}$  est aussi un chemin de flux pour  $W'$ . Par induction, il existe donc un flux correspondant  $v = u; e \in W'$ . Si ce n'est pas le flux supprimé par la règle (DÉPENDANT), alors il existe un flux  $u; e' \in W$ , correspondant à  $\tilde{u}$ . On va alors montrer que  $v$  ne peut pas être le flux qui a été supprimé. Plus précisément, on suppose que c'est le flux supprimé, et on va montrer que cela contredit l'hypothèse  $\tilde{u}(X) \leq 1$  pour toute espèce intermédiaire  $X$ .

On suppose donc que  $v$  dépend de flux  $v_1, \dots, v_k \in W'$ , avec  $v_i = u_i; e_i$ . Alors, par induction, il existe toujours au moins un flux  $v_j$  tel que  $\text{Prod}_{\mathcal{G}}(v_j \mathbf{r}) = \text{Reac}_{\mathcal{G}}(v_j \mathbf{r}) = \emptyset$ . De plus, il existe un chemin de flux circulaire  $\tilde{u}' = u'_1 \dots u'_l$  correspondant à  $v_j$  dans  $W'$ . Il existe alors des espèces intermédiaires  $X_1, \dots, X_l$  telles que pour tout  $1 \leq i < l$ ,  $\text{Prod}_{\mathcal{G}}(u'_i \mathbf{r}) = \text{Reac}_{\mathcal{G}}(u'_{i+1} \mathbf{r}) = X_i$  et  $\text{Prod}_{\mathcal{G}}(u'_l \mathbf{r}) = \text{Reac}_{\mathcal{G}}(u'_1 \mathbf{r}) = X_l$ . Comme  $\sum \tilde{u} = u = \sum_{1 \leq i \leq k} u_i$ , et  $u_j = \sum \tilde{u}' = \sum_{1 \leq i \leq l} u'_i$ , les vecteurs unitaires  $u'_i$  apparaissent aussi dans le chemin de flux  $\tilde{u}$ . Donc chaque  $X_i$  est produit au moins une fois dans  $\tilde{u}$ . De plus, il existe au moins un  $X_i$  qui est produit en dehors des vecteurs venant de  $u_j$ , pour que  $\tilde{u}$  soit bien un chemin de flux, comme par exemple  $X_l$  dans le chemin ci-dessous :

$$\tilde{u} : \dots \rightarrow \overbrace{X_l \rightarrow X_1 \rightarrow \dots \rightarrow X_{l-1} \rightarrow X_l}^{\tilde{u}_j} \rightarrow \dots$$

Donc il existe un  $X_i$  avec  $\tilde{u}(X_i) > 1$ , ce qui contredit notre hypothèse initiale.

- (3) Soit  $v = u; e \in W$  un flux dépendant de  $v_1 = u_1; e_1, \dots, v_k = u_k; e_k \in W$  avec les coefficients  $a_1, \dots, a_k$ . Alors il existe aussi des flux  $v' = u; e'; v'_1 = u_1; e'_1, \dots, v'_k = u_k; e'_k \in W'$ , avec  $v'$  qui dépend de  $v'_1, \dots, v'_k$ . Par induction, les conditions du point (3) sont vérifiées dans  $W'$ , et directement préservées dans  $W$ .

On considère maintenant la règle (**INTERMÉDIAIRE**), éliminant l'espèce intermédiaire  $X$ , et on prouve le lemme point par point.

(1) Soit  $v = u; e \in W$  un flux. Il est soit la simplification d'un flux  $u; e' \in W'$ , et on peut directement conclure par induction. Ou il est la combinaison d'un flux  $v_p \in W'$  qui produit  $X$ , et d'un flux  $v_c \in WW'$  qui consomme  $X$ . Dans ce cas, par induction, il existe des chemins de flux correspondants  $\tilde{u}_p$  et  $\tilde{u}_c$ . On a alors  $Prod_{\mathcal{J}}(\tilde{u}_p \mathbf{r}) = X = Reac_{\mathcal{J}}(\tilde{u}_c \mathbf{r})$ . Donc la concaténation  $\tilde{u} = \tilde{u}_p \tilde{u}_c$  est un chemin; C'est de plus un chemin de flux :

- si  $\tilde{u}$  est non-circulaire, alors pour toute espèce intermédiaire  $Y$  telle que  $\tilde{u}(Y) > 0$ , soit  $Y = X$  et vient d'être éliminée, soit  $\tilde{u}_p(Y) > 0$  ou  $\tilde{u}_c(Y) > 0$ , et par induction  $Y$  a déjà été éliminée par la simplification dans  $W'$ , et donc dans  $W$ . Donc  $\tilde{u}$  est un chemin de flux non-circulaire.
- si  $\tilde{u}$  est circulaire, il existe une espèce intermédiaire  $Z$  consommée par  $v_p$  et produite par  $v_c$ . Le chemin de flux  $\tilde{u}_p$  ne peut pas être circulaire, sinon on aurait  $Reac_{\mathcal{J}}(u_p \mathbf{r}) = Prod_{\mathcal{J}}(u_p \mathbf{r}) = \emptyset \neq X$ , et de même pour  $\tilde{u}_c$ . Alors, comme précédemment, pour tout  $Y \neq Z$  est soit  $X$ , soit par induction une espèce qui a déjà été supprimée dans  $W'$ . Donc  $\tilde{u}$  est bien un chemin de flux circulaire.

De plus,  $\tilde{u}$  correspond trivialement à  $v$ . On prouve plus loin, en même temps que le point (3), que si  $\tilde{u}$  est non dépendant, alors  $\tilde{u}(X) \leq 1$  pour tout  $X$ .

(2) Soit  $\tilde{u}$  un chemin de flux pour  $W$  tel que pour toute espèce intermédiaire  $Y$ ,  $\tilde{u}(Y) \leq 1$ , et en particulier  $\tilde{u}(X) \leq 1$ , c'est-à-dire soit  $\tilde{u}(X) = 0$  ou  $\tilde{u}(X) = 1$ . Si  $\tilde{u}(X) = 0$ , comme  $Espèces(W) = Espèces(W') \setminus \{X\}$ ,  $\tilde{u}$  est aussi un chemin pour  $W'$ . Alors par induction, il existe un flux correspondant  $v = u; e \in W'$ . Comme  $\tilde{u}$  est un chemin de flux dans  $W$ , on sait aussi que  $X \notin Reac_{\mathcal{J}}(\tilde{u} \mathbf{r}) \cup Prod_{\mathcal{J}}(\tilde{u} \mathbf{r})$ . Donc  $X \notin Reac_{\mathcal{J}}(v \mathbf{r}) \cup Prod_{\mathcal{J}}(v \mathbf{r})$ . Il y a alors un flux  $v' = u; e' \in W$ , qui alors correspond à  $\tilde{u}$ . Si  $\tilde{u} = 1$ , alors on peut décomposer ce chemin en deux, un chemin  $\tilde{u}_p$  tel que  $Prod_{\mathcal{J}}(\tilde{u}_p \mathbf{r}) = X$ , et un chemin  $\tilde{u}_c$  tel que  $Reac_{\mathcal{J}}(\tilde{u}_c \mathbf{r}) = X$ , avec  $\tilde{u} = \tilde{u}_p \tilde{u}_c$ . Ces deux chemins de flux ne peuvent pas être circulaires, car cela impliquerait que  $\tilde{u}$  serait aussi circulaire et que  $\tilde{u}(X) > 1$ , contrairement à notre hypothèse. On a donc  $\tilde{u}_p(X) = \tilde{u}_c(X) = 0$ . Donc, comme  $Espèces(W) = Espèces(W') \setminus \{X\}$ , toutes les espèces  $Y$  telles que  $\tilde{u}_p(Y) \geq 1$  a déjà été supprimée dans  $W'$  (et pareil pour  $\tilde{u}_c$ ). Par induction, on a donc des flux correspondants  $v_p$  et  $v_c$  dans  $W'$ , le premier produisant  $X$  et le second le consommant. Il existe alors un flux  $v_p \diamond v_c$  dans  $W$ , qui correspond au chemin de flux  $\tilde{u}$ .

(3) On va aussi prouver ici la deuxième partie du point (1). Soit  $v = u; e \in W$  un flux. On distingue deux cas : soit  $v$  est la simplification d'un flux  $v' = u; e' \in$

$W'$  (donc  $v'$  ne produit pas et ni ne consomme  $X$ ), soit c'est la combinaison d'un flux  $v_p$  produisant  $X$  et d'un flux  $v_c$  le consommant.

On considère tout d'abord le premier cas. Par induction, et par le point (1), il existe un chemin de flux  $\tilde{u}$  pour  $W'$  correspondant à  $v'$ , et  $\tilde{u}(X) = 0$ , puisque  $X$  n'a pas encore été supprimée dans  $W'$ . Supposons tout d'abord que  $v$  dépend de  $v_1, \dots, v_k \in W$ , et que l'un des flux  $v_i$  est la combinaison de flux de  $W'$  produisant et consommant  $X$ . Dans ce cas, pour tout chemin  $\tilde{u}_i$  correspondant à  $v_i$ , on a  $\tilde{u}_i(X) > 0$ . Mais cela implique alors  $\tilde{u}(X) > 0$ , ce qui contredit notre hypothèse. Donc tous les  $v_i = u_i; e_i$  sont la simplification de flux  $v'_i = u_i; e'_i \in W'$ . Donc  $v'$  est dépendant de ces flux dans  $W'$ . Donc, par induction, le point (3) est vérifié pour  $v'$  dans  $W'$ , et donc aussi pour  $v$  dans  $W$ .

Maintenant, supposons que  $v$  ne soit pas dépendant dans  $W$ . Avec le même raisonnement que ci-dessus,  $\tilde{u}(X) = 0$ , donc si  $v'$  est dépendant de flux  $v'_i = u'_i; e'_i \in W'$ , ceux-ci ne produisent pas ni ne consomment  $X$ . Ils seraient donc simplifiés par (INTERMÉDIAIRE) en des flux  $v_i = u'_i; e_i \in W$ , et  $v$  dépendrait des  $v_i$ , ce qui est contradictoire. Donc  $v'$  n'est pas dépendant, donc par induction la deuxième partie du point (1) est vérifiée pour  $v'$ , et donc pour  $v$ , c'est-à-dire que pour toute espèce intermédiaire  $y$ , on a  $\tilde{u}(Y) \leq 1$ .

On considère maintenant le deuxième cas, où  $v$  est la combinaison de  $v_p = u_p; e$  et  $v_c = u_c; e_c \in W'$ . Par induction, il existe des chemins  $\tilde{u}_p$  et  $\tilde{u}_c$  correspondant à ces flux. Tout chemin  $\tilde{u}$  correspondant à  $v$  est alors la concaténation de  $\tilde{u}_p$  et  $\tilde{u}_c$ .

On commence par prouver que si  $v_p$  (ou  $v_c$ , ou les deux) est dépendant, alors  $v$  aussi. Supposons que  $v_p$  dépend de  $v'_1, \dots, v'_k \in W'$ . On sait que  $v_p$  produit du  $X$ , donc  $Prod_{\mathcal{J}}(v_p \mathbf{r}) = \{X\} \neq \emptyset$ . Par induction, il existe un (unique) index  $i$  tel que  $v'_i$  produise aussi du  $X$ , et que pour tout  $j \neq i$ ,  $Reac_{\mathcal{J}}(v'_j \mathbf{r}) = Prod_{\mathcal{J}}(v'_j \mathbf{r}) = \emptyset$ . Donc tout les flux  $v'_j = u'_j; e'_j$  vont être simplifiés en flux  $v_j = u'_j; e_j \in W$ , tandis que  $v'_i$  sera combiné avec les flux consommant  $X$ , et en particulier avec  $v_c$ . Dans ce cas,  $v$  va directement dépendre des  $v_j$  et de  $v'_i \diamond v_c$ .

On considère maintenant le cas où  $v_p$  et  $v_c$  sont tous les deux non dépendent. Par induction, pour toute espèce intermédiaire  $Y \neq X$ , on a donc  $\tilde{u}_p(Y) \leq 1$  et  $\tilde{u}_c(Y) \leq 1$ . Si  $\tilde{u}_p(Y) + \tilde{u}_c(Y) \leq 1$ , on a directement  $\tilde{u}(Y) \leq 1$ . De plus,  $\tilde{u}(X) = 1 \leq 1$  (en effet,  $X$  étant l'espèce supprimée par (INTERMÉDIAIRE), on a  $\tilde{u}_p(X) = \tilde{u}_c(X) = 0$ , et le  $X$  produit par  $\tilde{u}_p(X)$  est fusionné avec celui consommé par  $\tilde{u}_c(X)$ ). Donc dans ce cas, la deuxième partie du point (1) est bien satisfaite. Supposons maintenant qu'il existe un  $Y$  tel que  $\tilde{u}_p(Y) = \tilde{u}_c(Y) = 1$ . Dans ce cas, il y a donc un cycle dans  $\tilde{u}$ . On considère alors un sous-cycle  $u_{\tilde{cyc}_1}$  de taille minimale dans  $\tilde{u}$ , et  $Z$  l'espèce intermédiaire en

bordure de ce cycle :

$$\tilde{u} : \underbrace{\dots \rightarrow Y \rightarrow \dots}_{\tilde{u}_1} \xrightarrow{\tilde{u}_p} \underbrace{Z \rightarrow \dots \rightarrow X \rightarrow \dots \rightarrow Z}_{u_{\tilde{cyc}_1}} \xrightarrow{\tilde{u}_c} \underbrace{\dots \rightarrow Y \rightarrow \dots}_{\tilde{u}_2}$$

On note  $\tilde{u}_1$  (resp.  $\tilde{u}_2$ ) la partie de  $\tilde{u}$  précédent (resp. suivant) le cycle  $u_{\tilde{cyc}_1}$ , et  $u_{\tilde{reste}_1} = \tilde{u}_1 \tilde{u}_2$  la concaténation de ces deux chemins, c'est-à-dire le chemin restant lorsque l'on supprime le cycle  $u_{\tilde{cyc}_1}$  du chemin  $\tilde{u}$ . Le chemin de flux  $u_{\tilde{cyc}_1}$  vérifie que pour tout  $Y$ ,  $u_{\tilde{cyc}_1}(Y) \leq 1$ . Donc, par le point (2), il existe un flux correspondant  $v_{cic_1}$  dans  $W'$ , avec de plus  $Reac_{\mathcal{J}}(v_{cic_1} \mathbf{r}) = Prod_{\mathcal{J}}(v_{cic_1} \mathbf{r}) = \emptyset$ . On répète alors la même opération, mais cette fois sur  $u_{\tilde{reste}_1}$ . À chaque étape, on obtient un nouveau flux circulaire  $v_{cic_i}$ . On arrête lorsque dans le chemin restant  $u_{\tilde{reste}_k}$ , pour tout  $Y$ , on ait  $u_{\tilde{reste}_k}(Y) \leq 1$ . Dans ce cas, toujours par le point (2), il existe un flux  $v_{reste}$  correspondant, avec  $Reac_{\mathcal{J}}(v_{reste} \mathbf{r}) = Reac_{\mathcal{J}}(v \mathbf{r})$  et  $Prod_{\mathcal{J}}(v_{reste} \mathbf{r}) = Prod_{\mathcal{J}}(v \mathbf{r})$ . Dans ce cas,  $v$  est dépendant de  $v_{reste}$  et de l'ensemble des  $v_{cyc_i}$ , et le point (3) est satisfait. Supposons maintenant que le cas où  $v_p$  et  $v_c$  sont tous les deux dépendants (notons que le cas où un seul des deux est dépendant est en fait similaire à ce cas là, et ne sera donc pas traité en détail). Le flux  $v_p$  dépend de  $v_{p,1}, \dots, v_{p,k_p}$ , et  $v_c$  dépend de  $v_{c,1}, \dots, v_{c,k_c}$ . Comme  $v_p$  produit  $X$  et que  $v_c$  le consomme, ces flux ne sont pas circulaires. Donc par induction, il existe un unique  $i_p$  tel que  $Prod_{\mathcal{J}}(v_{p,i_p} \mathbf{r}) = Prod_{\mathcal{J}}(v \mathbf{r}) = \{X\}$ ,  $Reac_{\mathcal{J}}(v_{p,i_p} \mathbf{r}) = Reac_{\mathcal{J}}(v \mathbf{r})$ , et pour tout flux  $j \neq i_p$ ,  $Prod_{\mathcal{J}}(v_{p,j} \mathbf{r}) = Reac_{\mathcal{J}}(v_{p,j} \mathbf{r}) = \emptyset$  (notons en particulier que cela implique que  $v_c$  n'est pas l'un des  $v_{p,i}$ ). Et il existe un  $i_c$  avec les propriétés similaires pour  $v_c$ . Alors pour tout flux  $v_{p,j} = u_{p,j}; e_{p,j}$  est transformé en un flux  $v'_{p,j} = u_{p,j}; e'_{p,j}$  dans  $W$ , et similairement pour les  $v_{c,j}$ , tandis que  $v_{p,i_p}$  est combiné avec  $v_{c,i_c}$ , formant un nouveau flux  $v_i$  dans  $W$ . Alors  $v$  dépend directement de  $v_i$ , des  $v'_{p,j}$  et des  $v'_{c,j}$ , et ces flux vérifient bien les conditions du point (3). Dans le cas où par exemple seul  $v_p$  est dépendant, le raisonnement est identique, en remplaçant  $v_{c,i_c}$  par directement  $v_c$ .

□

La Prop. 16 est alors une direct conséquence du Lem. 13.

### 5.2.6 Stabilité de *LinNets*

On peut alors utiliser la Prop. 16 pour prouver que l'ensemble *LinNets* est stable par simplification.

**Proposition 17.** Soit  $W \in \text{LinNets}$  et  $W \Rightarrow W'$ . Alors  $W' \in \text{LinNets}$ .



*Proof.* Supposons qu'on simplifie  $W$  en  $W'$  avec la règle (DÉPENDANT). Comme  $W \in \text{LinNets}$ , la condition 4 sur la stœchiométrie est vérifiée dans  $W$ , et donc également dans  $W'$ .

Soit  $X$  une espèce intermédiaire, et  $v_d = u_d; e_d \in W$  le flux supprimé, dépendant de  $v_1, \dots, v_k \in W$ , avec des coefficients  $a_1, \dots, a_k \in \mathbb{N}$ . Par induction, dans  $W$  l'espèce  $X$  est soit un modificateur, soit est stable-linéaire. Si  $X$  est un modificateur, c'est aussi le cas dans  $W'$ , et toutes les autres conditions sont directement vérifiées.

Sinon, soit  $v = u; e$  un flux de  $W'$ . Si  $v \in W$ , c'est-à-dire que ce flux n'a pas été modifié par la règle, alors les conditions sur sa cinétique sont vérifiées dans  $W$  par induction, et donc également dans  $W'$ . Sinon, cela signifie qu'il existe  $i$  tel que  $v$  soit la simplification du flux  $v_i = u_i; e_i$ . Dans ce cas, on a donc  $u = u_i$ , et  $e = e_i + a_i e_d$ . D'après la Prop. 16, soit  $\text{Prod}_{\mathcal{J}}(v_i \mathbf{r}) = \text{Reac}_{\mathcal{J}}(v_i \mathbf{r}) = \emptyset$ , soit  $\text{Reac}_{\mathcal{J}}(v_i \mathbf{r}) = \text{Reac}_{\mathcal{J}}(v_d \mathbf{r})$  et  $\text{Prod}_{\mathcal{J}}(v_i \mathbf{r}) = \text{Prod}_{\mathcal{J}}(v_d \mathbf{r})$ . Comme  $u_i = u$ , c'est également le cas en remplaçant  $v_i$  par  $v$ . Dans le premier cas,  $v$  ne produit ni ne consomme aucune espèce intermédiaire, donc les conditions sur les cinétiques sont trivialement vérifiées. Dans le second cas, supposons que  $v$  consomme  $X$ . Dans ce cas, c'est aussi le cas de  $v_i$  et de  $v_d$ . Donc par induction, comme  $W \in \text{LinNets}$ , on a  $e_d = x_X e'_d$  et  $e_i = x_X e'_i$ , avec  $X \notin \text{Espèces}(e'_d), \text{Espèces}(e'_i)$ , et pour toute espèce intermédiaire  $Y \neq X$ ,  $Y \notin \text{Espèces}(e_d), \text{Espèces}(e_i)$ . Donc  $e = e_i + a_i e_d = x_X (e'_i + a_i e'_d)$  est bien linéaire en  $X$ , et pour tout  $Y \neq X$ ,  $Y \notin \text{Espèces}(e)$ . Donc le point 2 de la définition de  $\text{LinNets}$  est vérifié.

Si  $v$  produit  $X$  et ne consomme pas d'autres espèces intermédiaires, alors c'est aussi le cas de  $v_i$  et  $v_d$ , donc par induction,  $\mathcal{J} \cap \text{Espèces}(e_i) = \emptyset$ , et pareil pour  $e_d$ . Comme  $e = e_i + a_i e_d$ , on a donc  $\mathcal{J} \cap \text{Espèces}(e) = \emptyset$ , donc le point 3 est également vérifié. De plus cela prouve également que  $X$  est bien stable-linéaire dans  $W'$  (point 1). Donc  $W' \in \text{LinNets}$ .

Si on simplifie  $W$  en  $W'$  en utilisant la règle (MODIFICATEUR CONSTANT), alors les conditions de  $\text{LinNets}$  sont trivialement préservées.

Enfin, si on utilise la règle (INTERMÉDIAIRE) sur une espèce  $X \in \mathcal{J}$ . Soient  $T, T'$  et  $F$  les expressions définies dans la règle. Les règles de linéarité dans  $W$  impliquent que pour tout  $Y \in \mathcal{J}$ ,  $x_Y \notin \text{Vars}(T')$ .

Soit  $Y \neq X$  une autre espèce intermédiaire. Soit  $Y$  est un modificateur dans  $W'$ , et donc toutes les conditions sont directement vérifiées. Soit ce n'est pas le cas, donc il existe un flux produisant ou consommant  $Y$ . Soit  $v = u; e \in W'$  un tel flux, et supposons que  $v$  consomme  $Y$ . Supposons tout d'abord que  $v$  soit la simplification d'un flux  $v' = u; e' \in W$ , c'est-à-dire  $e = e'[x_X := F]$ . Par le point 2, à part  $Y$ , aucune autre espèce intermédiaire ne peut apparaître dans  $e'$ . En particulier,  $X \notin \text{Espèces}(e')$ , et donc  $e = e'$ . Donc le point 2 est directement vérifié

pour  $v$  dans ce cas. Si  $v$  est la combinaison d'un flux  $v_p$  qui produit  $X$  et d'un flux  $v_c$  qui consomme  $X$ , on a alors  $v = v_p \diamond_{T'} v_c$ . On a  $Prod_{\mathcal{J}}(v_p \mathbf{r}) = Reac_{\mathcal{J}}(v_c \mathbf{r}) = \{X\}$ ,  $Reac_{\mathcal{J}}(v_p \mathbf{r}) = Reac_{\mathcal{J}}(v \mathbf{r}) = \{Y\}$ , et  $Prod_{\mathcal{J}}(v_c \mathbf{r}) = Prod_{\mathcal{J}}(v \mathbf{r})$ . Donc par induction,  $kin(v_p) = x_Y e_p$  ( $v_p$  consomme  $Y$ ) et  $kin(v_c) = x_X e_c$  ( $v_c$  consomme  $X$ ), avec pour tout  $Z \in \mathcal{J}$ ,  $Z \notin Espèces(e_c), Espèces(e_p)$ . Donc  $e = x_Y e_p e_c / T'$ , avec  $Z \notin Espèces(e_p e_c / T')$ , et donc la condition 2 est bien vérifiée.

Supposons maintenant que  $v = u; e \in W'$  produise  $Y$  et ne consomme pas d'espèce intermédiaire. Si c'est la simplification d'un  $v' = u; e' \in W$ , alors comme précédemment, on a  $X \notin Espèces(e')$  par induction et par le point 3, et donc  $e = e'$  et toutes les conditions sur les cinétiques sont vérifiées. Si  $v'$  est la fusion d'un  $v_p$  et d'un  $v_c$ , alors  $Reac_{\mathcal{J}}(v_p \mathbf{r}) = \emptyset$ ,  $Prod_{\mathcal{J}}(v_p \mathbf{r}) = Reac_{\mathcal{J}}(v_c \mathbf{r}) = \{X\}$ , et  $Prod_{\mathcal{J}}(v_c \mathbf{r}) = \{Y\}$ . Donc  $\mathcal{J} \cap Espèces(kin(v_p)) = \emptyset$ , et  $kin(v_c) = x_X e'_c$ , avec  $\mathcal{J} \cap Espèces(e'_c) = \emptyset$ . Donc comme  $kin(v) = kin(v_p) e'_c / T'$ , on a  $\mathcal{J} \cap Espèces(kin(v)) = \emptyset$ , et le point 3 est vérifié.

Cela prouve également le point 1.

Pour la condition 4 sur la stoechiométrie, pour tout flux  $v = v_p \diamond_{T'} v_c$ , on a avec la normalisation :

$$\begin{aligned} Reac_{\mathcal{J}}(v \mathbf{r}) &= (Reac_{\mathcal{J}}(v_p \mathbf{r}) \cup Reac_{\mathcal{J}}(v_c \mathbf{r})) \setminus (Prod_{\mathcal{J}}(v_p \mathbf{r}) \cup Prod_{\mathcal{J}}(v_c \mathbf{r})) \\ &= (Reac_{\mathcal{J}}(v_p \mathbf{r}) \cup \{X\}) \setminus (\{X\} \cup Prod_{\mathcal{J}}(v_c \mathbf{r})) \\ &= Reac_{\mathcal{J}}(v_p \mathbf{r}) \setminus (\{X\} \cup Prod_{\mathcal{J}}(v_c \mathbf{r})). \end{aligned}$$

Comme  $|Reac_{\mathcal{J}}(v_p \mathbf{r})| \leq 1$ , on a directement  $|Reac_{\mathcal{J}}(v \mathbf{r})| \leq 1$ . Par symétrie,  $|Prod_{\mathcal{J}}(v \mathbf{r})| \leq 1$ .

Donc  $W' \in LinNets$ . □

### 5.2.7 Confluence structurelle

On peut maintenant prouver que la simplification est structurellement confluente, c'est-à-dire que tous les réseaux complètement simplifiés auront les mêmes réactions si on néglige les cinétiques.

On définit la taille d'un réseau de flux  $W$  comme celle du réseau de réactions  $W \mathbf{r}$  correspondant. Nos règles de simplification diminuent trivialement la taille des réseaux, donc la confluence locale implique la confluence globale.

**Définition 26.** Deux réseaux de flux  $W = V \& C$  et  $W' = V' \& C'$  sont *structurellement similaires*, noté  $W \equiv^{struc} W'$ , s'ils ont les mêmes réactions sans cinétiques, c'est-à-dire :

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

On a alors le théorème de confluence structurelle suivant.

**Théorème 7** (Structural confluence). La simplification  $\Rightarrow$  sur  $(LinNets, \equiv^{struc})$  est confluente.

*Proof.* La proposition 17 implique qu'on peut toujours simplifier toutes les espèces intermédiaires. Donc tout réseau complètement simplifié n'a plus d'espèces intermédiaires, ni d'espèces dépendantes. Le Théorème 6 implique alors l'unicité de la structure.  $\square$

## 5.3 Confluence générale de la simplification avec cinétique

Enfin, on s'intéresse à la confluence générale de la simplification avec cinétique. On veut donc que les réseaux complètement simplifiés aient la même structure, mais aussi les mêmes expressions cinétiques et les mêmes contraintes. On montre tout d'abord sur un exemple que ce n'est pas le cas en général. On donne alors un critère sur les modes du réseau initial, et on prouve que lorsque ce critère est vérifié, la simplification est confluente. On donne ensuite un exemple biologique de modèle dans lequel le critère n'est pas vérifié, et où on peut obtenir deux réseaux simplifiés différents. Enfin, on étudie la simplification des systèmes d'équations qu'on peut dériver des réseaux.

### 5.3.1 Non confluence des expressions cinétiques

L'exemple suivant illustre la non confluence des expressions cinétiques dans le cas général.

**Exemple 52.** Reprenons le réseau  $W$  de la Fig. 5.6. On peut éliminer les espèces intermédiaires avec la règle (**INTERMÉDIAIRE**). Si on commence par éliminer  $X$ , puis  $Y$  et enfin  $Z$ , on obtient le réseau de flux  $W_{XYZ}$  représenté en haut à droite de la Fig. 5.6. Si on commence par  $X$ , puis  $Z$  puis  $Y$ , on obtient le réseau  $W_{XZY}$  en bas à droite. Dans le réseau  $W_{XYZ}$ , le flux  $v_{123456}$  est dépendant, soit des flux  $v_{123}$ ,  $v_{456}$ , soit des flux  $v_{1346}$ ,  $v_{25}$ . On peut donc le supprimer de deux façon différentes avec la règle (**DÉPENDANT**). Dans le premier cas, on modifie alors les cinétiques des flux  $v_{123}$  et  $v_{456}$ , et on obtient le réseau  $W_{XYZd_1} = W_{XZY}$ . Dans le second cas, on modifie les cinétiques de  $v_{1346}$  et  $v_{25}$ , et on obtient le réseau  $W_{XYZd_2}$  (Fig. 5.6).

Ces deux réseaux ont la même structure, c'est-à-dire les mêmes réactions lorsqu'on néglige les cinétiques. En revanche, les expressions cinétiques des flux ne sont pas les mêmes.

### 5.3.2 Critère pour la confluence

On peut alors définir le critère suivant sur la base de réactions initiales.

**Définition 27.** Une base de réactions  $\mathbf{r}$  est *décomposable de façon unique* si tout mode  $u \in \mathbb{N}^m$  pour les espèces intermédiaires  $\mathcal{J}$  a une unique décomposition en modes élémentaires.

**Exemple 53.** Dans l'exemple précédent, la base de réaction pour le réseau  $W$  (Fig. 5.6) n'est pas uniquement décomposable. Il a quatre modes élémentaires :

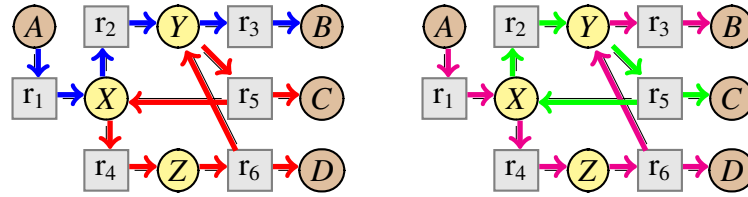


Fig 5.9: Deux décompositions du mode  $u = (1, 1, 1, 1, 1, 1)$  dans le réseau  $W$ .

$$u_1 = (1, 1, 1, 0, 0, 0) \quad u_2 = (0, 0, 0, 1, 1, 1)$$

$$u_3 = (1, 0, 1, 1, 0, 1) \quad u_4 = (0, 1, 0, 0, 1, 0)$$

Alors le mode  $u = (1, 1, 1, 1, 1, 1)$  possède deux décompositions différentes :  $u_1 + u_2$  et  $u_3 + u_4$ . Ces deux décompositions sont illustrées dans la Fig. 5.9. Le critère n'est donc pas vérifié, et la simplification n'est pas confluente, comme on l'a vu précédemment.

### 5.3.3 Confluence

On prouve alors que lorsque le critère de décomposition est respecté, la simplification est confluente, y compris pour les cinétiques.

**Théorème 8** (Confluence). Si  $\mathbf{r}$  est décomposable de façon unique, alors la simplification  $\Rightarrow$  sur  $(LinNets, \equiv)$  est confluente.

*Proof.* C'est la conséquence des lemmes 14 à 19 ci-dessous, qui analysent les différentes paires critiques. Notons que puisque qu'on considère la confluence modulo la similarité  $\equiv$ , la confluence de la règle (SIMILARITÉ) avec les autres règles est triviale.  $\square$

Les lemmes suivants analysent les différentes paires critiques. En particulier, les preuves des lemmes 16 et 19 nécessitent de décomposer précisément les différents ensembles de flux.

**Lemme 14.** On suppose que la base  $\mathbf{r}$  est décomposable de façon unique. Soit  $W$  un réseau tel qu'on puisse le simplifier en  $W_1$  ou  $W_2$  avec la règle (DÉPENDANT) :  $W \Rightarrow^{Dep} W_i$ . Alors il existe deux réseaux similaires  $W'_1 \equiv W'_2$  tels que  $W_i \Rightarrow^* W'_i$ .

*Proof.* Supposons qu'on simplifie  $W$  en  $W_i$  en appliquant la règle (DÉPENDANT) sur respectivement  $v_d$  qui dépend de  $v_1, \dots, v_k$ , et  $v'_d$ , qui dépend de  $v'_1, \dots, v'_k$ . Notons que dans tous les cas, on ne modifie pas la contrainte de  $W$ .

Si  $v_d = v'_d$ , alors comme  $\mathbf{r}$  est décomposable de façon unique, cela implique  $\{v_1, \dots, v_k\} = \{v'_1, \dots, v'_k\}$ , et donc directement  $W_1 = W_2$ .

Si  $v_d \neq v'_d$ , que  $v_d \notin \{v'_1, \dots, v'_k\}$ , et  $v \notin \{v_1, \dots, v_k\}$ , alors la paire critique se ferme uniformément (modulo la similarité des cinétiques) :  $W_i \Rightarrow^{Dep} W'_i$ , avec  $W'_1 \equiv W'_2$ .

Si  $v_d$  dépend de  $v'_d$ , alors  $v'_d$  ne peut pas dépendre de  $v_d$ , puisqu'on utilise uniquement des combinaisons linéaires positives. On suppose, sans perdre en généralité, que  $v'_d = v_1$ . On note  $v_d = u_d; e_d$ ,  $v'_d = u'_d; e'_d$ ,  $v_i = u_i; e_i$  et  $v'_i = u'_i; e'_i$ . Il existe des entiers positifs  $a'_d$ ,  $a_i$  et  $a'_i$  tels que  $u_d = a'_d u'_d + \sum_{2 \leq i \leq k} a_i u_i$ , et

$u'_d = \sum_{1 \leq i \leq k'} a'_i u'_i$ . Dans  $W_1$ , le flux  $v_d$  est supprimé, et le flux  $v'_d$  devient  $u'_d; e'_d + a'_d e_d$ .

Il est toujours dépendant, donc on peut le supprimer. Il reste donc des flux de forme  $u_i; e_i + a_i e_d$  et  $u'_i; e'_i + a'_i (e'_d + a'_d e_d)$ . Dans  $W_2$ , après avoir supprimé  $v'_d$ , le flux  $v_d$  reste dépendant de  $\{v'_1, \dots, v'_k, v_2, \dots, v_k\}$ , avec les coefficients suivants :  $u_d = \sum_{1 \leq i \leq k'} a'_d a'_i u'_i + \sum_{2 \leq i \leq k} a_i u_i$ . On peut donc le supprimer. Il reste des flux de forme  $u_i; e_i + a_i e_d$  et  $u'_i; e'_i + a'_i e'_d + a'_d a'_i e_d$ , d'où  $W'_1 \equiv W'_2$ .  $\square$

**Lemme 15.** Soit  $W$  un réseau tel qu'on puisse le simplifier en  $W_1$  avec la règle (DÉPENDANT),  $W \Rightarrow^{Dep} W_1$ , ou en  $W_2$  avec la règle (MODIFICATEUR CONSTANT),  $W \Rightarrow^{Mod} W_2$ . Alors  $W_1 \Rightarrow^{Mod} W'_2$  et  $W_2 \Rightarrow^{Dep} W'_1$ , avec  $W'_1 \equiv W'_2$ .

*Proof.* Ce cas est direct. Soit  $v_d = u_d; e_d$  le flux supprimé,  $X$  le modificateur, et  $v = u; e \in W$  un autre flux. On suppose que  $v_d$  dépend de  $v$  avec un coefficient  $a$  (et avec  $a = 0$  si  $v_d$  ne dépend pas de  $v$ ). Dans  $W'_1$ , ce flux devient  $u; (e + a_d e_d)[x_X := x_X^0]$ . Dans  $W'_2$ , il devient  $u; e[x_X := x_X^0] + a e_d[x_X := x_X^0]$ . Ces deux flux sont trivialement similaires. De plus, dans les deux cas on obtient la contrainte  $C[x_X := x_X^0]$ , donc  $W'_1 \equiv W'_2$ .  $\square$

**Lemme 16.** Soit  $W$  un réseau tel qu'on puisse le simplifier en  $W_1$  avec la règle (DÉPENDANT),  $W \Rightarrow^{Dep} W_1$ , ou en  $W_2$  avec la règle (INTERMÉDIAIRE),  $W \Rightarrow^{Inter} W_2$ . Alors il existe  $W'_1 \equiv W'_2$  avec  $W_i \Rightarrow^* W'_i$ .

*Proof.* Cette preuve est un peu plus longue. L'idée est que, en utilisant la décomposition de la Prop. 16, si le flux dépendant  $v_d$  par exemple consomme l'espèce intermédiaire  $X$ , alors c'est aussi le cas pour un (unique) flux  $v_i$  dont  $v_d$  dépend. Dans ce cas, pour tout flux  $v$  consommant  $X$ , le flux combiné  $v_d \diamond v$  sera toujours dépendant dans  $W_2$ , avec notamment le flux  $v_i \diamond v$ .

Soit  $v_d$  le flux dépendant supprimé, qui dépend de  $v_1, \dots, v_k$ , avec les coefficients  $a_1, \dots, a_k$ .

Considérons tout d'abord que  $X \notin \text{Prod}_{\mathcal{J}}(v_d \mathbf{r}) \cup \text{Reac}_{\mathcal{J}}(v_d \mathbf{r})$ . On note  $v_d = u_d; e_d$  et pour tout  $1 \leq i \leq k$ ,  $v_i = u_i; e_i$ . Alors, d'après la Prop. 16, pour tout  $1 \leq$

$i \leq k$ ,  $X \notin \text{Prod}_{\mathcal{G}}(v_i \mathbf{r}) \cup \text{Reac}_{\mathcal{G}}(v_i \mathbf{r})$ . Donc, après avoir éliminer l'espèce intermédiaire  $X$  dans  $W_2$ , on obtient les flux  $u_d; e_d[x_X := F], u_1; e_1[x_X := F], \dots, u_k; e_k[x_X := F] \in W_2$ , tels que  $u_d; e_d[x_X := F]$  dépend des autres flux. Avec  $F$  défini comme dans la règle (**INTERMÉDIAIRE**). On peut donc supprimer  $u_d; e_d[x_X := F]$  avec la règle (**DÉPENDANT**), et obtenir dans  $W'_2$  les flux  $u_i; e_i[x_X := F] + a_i e_d[x_X := F]$  pour  $1 \leq i \leq k$ . Réciproquement, en commençant par la règle (**DÉPENDANT**), on obtient dans  $W_1$  les flux  $u_i; e_i + a_i e_d$ . On peut alors toujours appliquer (**INTERMÉDIAIRE**), et on obtient alors  $u_i; (e_i + a_i e_d)[x_X := F]$  dans  $W'_1$ . Ces flux sont similaires à ceux de  $W'_1$ . C'est également le cas des autres flux de ces réseaux, et de leurs contraintes, donc  $W'_1 \equiv W'_2$ .

Supposons maintenant que  $v_d$  consomme  $X$ . Alors, d'après la Prop. 16, il existe un index  $i$  tel que  $v_i$  consomme  $X$  et  $a_i = 1$ , et pour tout  $j \neq i$ ,  $\text{Prod}_{\mathcal{G}}(v_i \mathbf{r}) = \text{Reac}_{\mathcal{G}}(v_i \mathbf{r}) = \emptyset$ . On note  $v_d = u_d; x_X e_d$ ,  $v_i = u_i; x_X e_i$  et pour tout  $j \neq i$ ,  $v_j = u_j; e_j$ . On note  $V_{prod}$  l'ensemble des flux qui produisent des molécules de  $X$ , et  $V_{reac}$  ceux qui en consomment. Dans  $W_2$ , après avoir supprimé l'espèce  $X$ , on obtient les flux suivants :

- $V_1$ , la combinaison de  $v_d$  et des flux consommant  $X$  :  $\{u + u_d; e_d e / T' \mid u; e \in V_{prod}\}$ ,
- $V_2$ , la combinaison de  $v_i$  et des flux consommant  $X$  :  $\{u + u_i; e_i e / T' \mid u; e \in V_{prod}\}$ ,
- $V_3$ , les autres flux combinés :  $\{u + u'; e e' / T' \mid u; e \in V_{prod}, u'; x_X e' \in V_{reac} \setminus \{v_d, v_i\}\}$ ,
- $V_4$ , les autres flux dont  $v_d$  dépend:  $\{u_j; e_j[x_X := F]\}_{j \neq i}$ ,
- $V_5$ , tous les autres flux, où on remplace juste  $x_X$  par  $F$ .

Alors tout flux  $u + u_d; e_d e / T' \in V_1$  dépend de  $u + u_i; e_i e / T' \in V_2$  et des flux  $u_j; e_j[x_X := F] \in VV_4$ , et peut donc être supprimé. On obtient les flux :

- $V'_2 = \{u + u_i; e_i e / T' + e_d e / T' \mid u; e \in V_{prod}\}$ ,
- $V_3$ ,
- $V'_4 = \{u_j; e_j[x_X := F] + \sum_{u; e \in V_{prod}} a_j e_d e / T'\}_{j \neq i}$ ,
- $V_5$ .

Réciproquement, en commençant par supprimer  $v_d$ , on obtient dans  $W_1$  :

- le flux  $v'_i = u_i; x_X e_i + x_X e_d$

- les autres flux dont  $v_d$  dépendait  $\{u_j; e_j + a_j x_X e_d\}_{j \neq i}$
- $V_{cons} \setminus \{v_d, v_i\}$ ,
- $V_{prod}$ ,
- les autres flux.

On peut toujours supprimer  $X$ , et obtenir :

- la combinaison de  $v'_i$  et des flux consommant  $X$  :  $\{u + u_i; (e_i + e_d)e/T' \mid u; e \in V_{prod}\} \equiv V'_2$ ,
- les autres combinaisons de flux :  $V_3$ ,
- les autres flux dont  $v_d$  dépendait :  $VV' = \{u_j; (e_j + a_j x_X e_d)[x_X := F]\}_{j \neq i}$ ,
- tous les autres flux, où on remplace juste  $x_X$  par  $F$  :  $V_5$ .

Or on a :

$$\begin{aligned} e_j[x_X := F] + \sum_{u; e \in V_{prod}} a_j e_d e/T' &\equiv e_j[x_X := F] + a_j e_d F \\ &\equiv (e_j + a_j x_X e_d)[x_X := F]. \end{aligned}$$

Donc  $V'' \equiv V'_4$ , et les flux des deux réseaux simplifiés sont similaires. Pour les contraintes, dans les deux réseaux on obtient  $C[x_X := F] \wedge x_X^0 = F$ .

Le cas où  $X$  est produit par  $v_d$  est similaire au précédent  $\square$

**Lemme 17.** Soit  $W$  un réseau tel qu'on puisse le simplifier en  $W_1$  ou  $W_2$  avec la règle (**MODIFICATEUR CONSTANT**) :  $W \Rightarrow^{Mod} W_i$ . Alors il existe  $W'_1 \equiv W'_2$  avec  $W_i \Rightarrow^{Mod} W'_i$ .

*Proof.* Ce cas est direct, puisque les substitutions commutent.  $\square$

**Lemme 18.** Soit  $W$  un réseau tel qu'on puisse le simplifier en  $W_1$  avec la règle (**MODIFICATEUR CONSTANT**),  $W \Rightarrow^{Mod} W_1$ , ou en  $W_2$  avec la règle (**INTERMÉDIAIRE**),  $W \Rightarrow^{Inter} W_2$ . Alors il existe  $W'_1 \equiv W'_2$  avec  $W_1 \Rightarrow^{Inter} W'_1$  et  $W_2 \Rightarrow^{Mod} W'_2$ .

*Proof.* Les espèces supprimées ne pouvant pas être les mêmes, ce cas est trivial.  $\square$

**Lemme 19.** Soit  $W$  un réseau tel qu'on puisse le simplifier en  $W_1$  ou  $W_2$  avec la règle (**INTERMÉDIAIRE**) :  $W \Rightarrow^{Inter} W_i$ . Alors il existe  $W'_1 \equiv W'_2$  avec  $W_i \Rightarrow^* W'_i$ .



*Proof.* Cette preuve est également plus longue. On va montrer qu'après avoir enlevé une première espèce intermédiaire  $X$ , l'autre espèce  $Y$  peut être supprimée soit avec (**MODIFICATEUR CONSTANT**), soit avec (**INTERMÉDIAIRE**). Dans le second cas, des flux dépendants peuvent éventuellement être créés, mais on peut les supprimer et obtenir le même réseau simplifié quelque soit l'ordre d'élimination de  $X$  et  $Y$ .

On partitionne les flux de  $W$  de la façon suivante :

- $V_X = \{v_X = u_X; e_X \mid X \in \text{Prod}_{\mathcal{G}}(v_X), Y \notin \text{Espèces}(v_X)\}$ , les flux produisant du  $X$  sans  $Y$ ,
- $V_{X'} = \{v_{X'} = u_{X'}; x_X e_{X'} \mid X \in \text{Reac}_{\mathcal{G}}(v_{X'}), Y \notin \text{Espèces}(v_{X'})\}$ , les flux consommant du  $X$  sans  $Y$ ,
- $V_{\text{mod}(X)} = \{v_{\text{mod}(X)} = u_{\text{mod}(X)}; e_{\text{mod}(X)} \mid X \notin \text{Prod}_{\mathcal{G}}(v_{\text{mod}(X)}) \cup \text{Reac}_{\mathcal{G}}(v_{\text{mod}(X)}), x_X \in \text{Vars}(v_{\text{mod}(X)}), Y \notin \text{Espèces}(v_{\text{mod}(X)})\}$ , les flux avec  $X$  comme modificateur et sans  $Y$ ,
- $V_Y = \{v_Y = u_Y; e_Y \mid Y \in \text{Prod}_{\mathcal{G}}(v_Y), X \notin \text{Espèces}(v_Y)\}$ , les flux produisant du  $Y$  sans  $X$ ,
- $V_{Y'} = \{v_{Y'} = u_{Y'}; x_Y e_{Y'} \mid Y \in \text{Reac}_{\mathcal{G}}(v_{Y'}), X \notin \text{Espèces}(v_{Y'})\}$ , les flux consommant du  $Y$  sans  $X$ ,
- $V_{\text{mod}(Y)} = \{v_{\text{mod}(Y)} = u_{\text{mod}(Y)}; e_{\text{mod}(Y)} \mid Y \notin \text{Prod}_{\mathcal{G}}(v_{\text{mod}(Y)}) \cup \text{Reac}_{\mathcal{G}}(v_{\text{mod}(Y)}), x_Y \in \text{Vars}(v_{\text{mod}(Y)}), X \notin \text{Espèces}(v_{\text{mod}(Y)})\}$ , les flux avec  $Y$  comme modificateur et sans  $X$ ,
- $V_{XY'} = \{v_{XY'} = u_{XY'}; x_Y e_{XY'} \mid X \in \text{Prod}_{\mathcal{G}}(v_{XY'}), Y \in \text{Reac}_{\mathcal{G}}(v_{XY'})\}$ , les flux produisant du  $X$  et consommant du  $Y$ ,
- $V_{X'Y} = \{v_{X'Y} = u_{X'Y}; x_X e_{X'Y} \mid Y \in \text{Prod}_{\mathcal{G}}(v_{X'Y}), X \in \text{Reac}_{\mathcal{G}}(v_{X'Y})\}$ , les flux produisant du  $Y$  et consommant du  $X$ ,
- $V_{\text{mod}(XY)} = \{v_{\text{mod}(XY)} = u_{\text{mod}(XY)}; e_{\text{mod}(XY)} \mid X, Y \notin \text{Prod}_{\mathcal{G}}(v_{\text{mod}(XY)}) \cup \text{Reac}_{\mathcal{G}}(v_{\text{mod}(XY)})\}$ , les flux avec  $X$  et  $Y$  en modificateur.

On définit les termes suivants :

$$\begin{aligned} T_X &= \sum_{V_X} e_X & T_{X'} &= \sum_{V_{X'}} e_{X'} \\ T_Y &= \sum_{V_Y} e_Y & T_{Y'} &= \sum_{V_{Y'}} e_{Y'} \\ T_{X'Y} &= \sum_{V_{X'Y}} e_{X'Y} & T_{XY'} &= \sum_{V_{XY'}} e_{XY'} \end{aligned}$$

On commence par supprimer  $X$ . On obtient les flux combinés suivants :

- $V_X \diamond_{(T_{X'}+T_{X'Y})} V_{X'} = \{u_X + u_{X'}; e_X e_{X'} / (T_{X'} + T_{X'Y})\},$
- $V_X \diamond_{(T_{X'}+T_{X'Y})} V_{X'Y} = \{u_X + u_{X'Y}; e_X e_{X'Y} / (T_{X'} + T_{X'Y})\},$
- $V_{XY'} \diamond_{(T_{X'}+T_{X'Y})} V_{X'} = \{u_{XY'} + u_{X'}; e_{XY'} e_{X'} x_Y / (T_{X'} + T_{X'Y})\},$
- $V_{XY'} \diamond_{(T_{X'}+T_{X'Y})} V_{X'Y} = \{u_{XY'} + u_{X'Y}; e_{XY'} e_{X'Y} x_Y / (T_{X'} + T_{X'Y})\}.$
- $V'_{mod(X)} = \{u_{mod(X)}; e_{mod(X)} [x_X := (T_X + x_Y T_{XY'}) / (T_{X'} + T_{X'Y})]\},$
- $V'_{mod(XY)} = \{u_{mod(XY)}; e_{mod(XY)} [x_X := (T_X + x_Y T_{XY'}) / (T_{X'} + T_{X'Y})]\}.$
- $V_Y = \{u_Y; e_Y\},$
- $V_{Y'} = \{u_Y; x_Y e_{Y'}\},$
- $V_{mod(Y)} = \{u_{mod(Y)}; e_{mod(Y)}\}.$

Il y a alors deux cas. Premièrement,  $Y$  peut maintenant être seulement un modificateur. Pour cela, il faut donc en particulier que  $V_Y = V_{Y'} = \emptyset$ . Comme on pouvait appliquer (**INTERMÉDIAIRE**) sur  $Y$  dans le réseau initial, on avait  $V_{XY'} \neq \emptyset$ . Mais alors il faut aussi avoir  $V_X = V_{X'} = \emptyset$ . En d'autres termes, tous les flux qui produisaient un  $X$  consommait un  $Y$ , et réciproquement. Dans ce cas, en supprimant  $X$ , il ne reste que les flux  $V_{XY'} \diamond_{(T_{X'}+T_{X'Y})} V_{X'Y}$ ,  $V'_{mod(X)}$ ,  $V_{mod(Y)}$  et  $V'_{mod(XY)}$ . En appliquant (**MODIFICATEUR CONSTANT**), on obtient alors les flux :

- $\{u_{XY'} + u_{X'Y}; e_{XY'} e_{X'Y} x_Y^0 / T_{X'Y}\}.$
- $\{u_{mod(X)}; e_{mod(X)} [x_X := x_Y^0 T_{XY'} / T_{X'Y}]\},$
- $\{u_{mod(Y)}; e_{mod(Y)} [x_Y := x_Y^0]\},$
- $\{u_{mod(XY)}; e_{mod(XY)} [x_X := x_Y^0 T_{XY'} / T_{X'Y}] [x_Y := x_Y^0]\}.$

On a également la contrainte  $x_X^0 = x_Y^0 T_{XY'} / T_{X'Y}$ , qui peut se réécrire en  $x_Y^0 = x_X^0 T_{X'Y} / T_{XY'}$ . En utilisant cette égalité pour réécrire les cinétiques précédentes, on trouve exactement (par symétrie) le réseau obtenu en supprimant d'abord  $Y$  puis le modificateur  $X$ .

Dans le second cas, on peut toujours supprimer  $Y$  avec la règle (**INTERMÉDIAIRE**). On définit les termes suivants :

$$U_Y = \frac{T_{X'} T_Y + T_Y T_{X'Y} + T_X T_{X'Y}}{T_{X'} + T_{X'Y}} \quad U_{Y'} = \frac{T_{X'} T_{Y'} + T_{Y'} T_{X'Y} + T_{X'} T_{XY'}}{T_{X'} + T_{X'Y}}$$

$$T = T_{X'} T_{Y'} + T_{Y'} T_{X'Y} + T_{X'} T_{XY'}.$$

On obtient alors les flux suivants :

- $(V_X \diamond_{(T_{X'}+T_{X'Y})} V_{X'Y}) \diamond_{U_{Y'}} (V_{XY'} \diamond_{(T_{X'}+T_{X'Y})} V_{X'}) = \{u_X + u_{X'Y} + u_{X'} + u_{XY'}; \frac{e_X e_{X'} e_{XY'} e_{X'Y}}{(T_{X'} + T_{X'Y})T}\},$
- $(V_X \diamond_{(T_{X'}+T_{X'Y})} V_{X'Y}) \diamond_{U_{Y'}} V_{Y'} = \{u_X + u_{X'Y} + u_{Y'}; \frac{e_X e_{Y'} e_{X'Y}}{T}\},$
- $V_Y \diamond_{U_{Y'}} (V_{XY'} \diamond_{(T_{X'}+T_{X'Y})} V_{X'}) = \{u_Y + u_{XY'} + u_{X'}; \frac{e_Y e_{XY'} e_{X'}}{T}\},$
- $V_Y \diamond_{U_{Y'}} V_{Y'} = \{u_Y + u_{Y'}; \frac{e_Y e_{Y'} (T_{X'} + T_{X'Y})}{T}\},$
- $(V_{XY'} \diamond_{(T_{X'}+T_{X'Y})} V_{X'Y})' = \{u_{XY'} + u_{X'Y}; \frac{T_{X'} T_Y + T_Y T_{X'Y} + T_X T_{X'Y}}{T(T_{X'} + T_{X'Y})} e_{XY'} e_{X'Y}\},$
- $V''_{mod(X)} = \{u_{mod(X)}; e_{mod(X)} [x_X := \frac{T_X T_{Y'} + T_X T_{XY'} + T_Y T_{XY'}}{T}]\},$
- $V''_{mod(XY)} = \{u_{mod(XY)}; e_{mod(XY)} [x_X := \frac{T_X T_{Y'} + T_X T_{XY'} + T_Y T_{XY'}}{T}$   
 $[x_Y := \frac{T_Y T_{X'} + T_Y T_{X'Y} + T_X T_{X'Y}}{T}]\},$
- $V'_{mod(Y)} = \{u_{mod(Y)}; e_{mod(Y)} [x_Y := \frac{T_Y T_{X'} + T_Y T_{X'Y} + T_X T_{X'Y}}{T}]\},$
- $V_X \diamond_{(T_{X'}+T_{X'Y})} V_{X'} = \{u_X + u_{X'}; e_X e_{X'} / (T_{X'} + T_{X'Y})\}$

On s'aperçoit alors que certains flux sont dépendants. Ainsi, chaque flux de  $(V_X \diamond_{(T_{X'}+T_{X'Y})} V_{X'Y}) \diamond_{U_{Y'}} (V_{XY'} \diamond_{(T_{X'}+T_{X'Y})} V_{X'})$  est en fait la somme d'un flux de  $(V_{XY'} \diamond_{(T_{X'}+T_{X'Y})} V_{X'Y})'$  et d'un flux de  $V_X \diamond_{(T_{X'}+T_{X'Y})} V_{X'}$ , et peut donc être supprimé. En réécrivant les expressions cinétiques par similarité, on obtient les flux suivants :

- $V_X \diamond_{(T_{X'}+T_{X'Y})} V_{X'} = \{u_X + u_{X'}; \frac{e_X e_{X'} (T_{Y'} + T_{X'Y})}{T}\},$
- $V_Y \diamond_{U_{Y'}} V_{Y'} = \{u_Y + u_{Y'}; \frac{e_Y e_{Y'} (T_{X'} + T_{X'Y})}{T}\},$
- $(V_X \diamond_{(T_{X'}+T_{X'Y})} V_{X'Y}) \diamond_{U_{Y'}} V_{Y'} = \{u_X + u_{X'Y} + u_{Y'}; \frac{e_X e_{Y'} e_{X'Y}}{T}\},$
- $V_Y \diamond_{U_{Y'}} (V_{XY'} \diamond_{(T_{X'}+T_{X'Y})} V_{X'}) = \{u_Y + u_{XY'} + u_{X'}; \frac{e_Y e_{XY'} e_{X'}}{T}\},$
- $V''_{mod(X)} = \{u_{mod(X)}; e_{mod(X)} [x_X := \frac{T_X T_{Y'} + T_X T_{XY'} + T_Y T_{XY'}}{T}]\},$
- $V'_{mod(Y)} = \{u_{mod(Y)}; e_{mod(Y)} [x_Y := \frac{T_Y T_{X'} + T_Y T_{X'Y} + T_X T_{X'Y}}{T}]\},$

- $(V_{XY'} \diamond_{(T_{X'}+T_{X'Y})} V_{X'Y})' = \{u_{XY'} + u_{X'Y}; \frac{(T_Y + T_X)e_{XY'}e_{X'Y}}{T}\},$
- $V''_{mod(XY)} = \{u_{mod(XY)}; e_{mod(XY)}[x_X := \frac{T_X T_{Y'} + T_X T_{X'Y} + T_Y T_{X'Y}}{T},$   
 $[x_Y := \frac{T_Y T_{X'} + T_Y T_{X'Y} + T_X T_{X'Y}}{T}]]\}.$

Mais on observe alors que  $X$  et  $Y$  sont en fait symétriques dans ces ensembles : le second ensemble correspond au premier en remplaçant  $X$  par  $Y$ , le quatrième au troisième, le sixième au cinquième. Dans les deux derniers ensembles,  $X$  et  $Y$  sont directement symétriques. Donc, si on avait commencé par  $Y$  puis  $X$ , on aurait obtenu le même réseau. On obtient également dans les deux cas la contrainte  $C$  en remplaçant  $x_X$  par  $\frac{T_X T_{Y'} + T_X T_{X'Y} + T_Y T_{X'Y}}{T}$  et  $x_Y$  par  $\frac{T_Y T_{X'} + T_Y T_{X'Y} + T_X T_{X'Y}}{T}$ .  $\square$

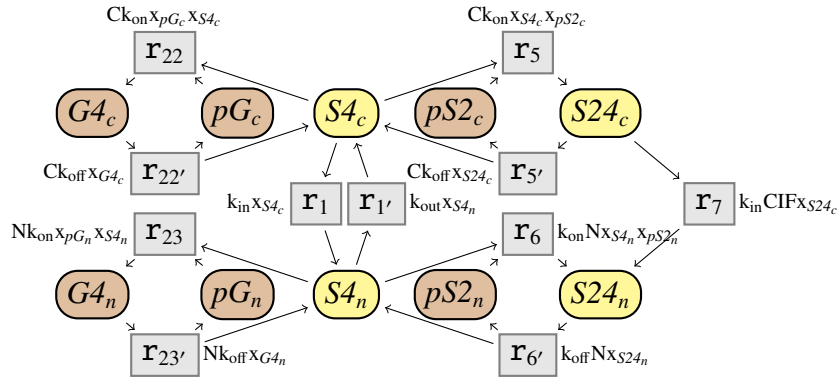
### 5.3.4 Exemple biologique de non confluence des cinétiques

On présente ici un exemple biologique d'un réseau ne vérifiant pas la condition de décomposition unique. On peut donc le simplifier de deux façons différentes, et obtenir deux réseaux avec la même structure, mais des fonctions cinétiques différentes.

Il s'agit du modèle *BIOMD0000000173* de la base de données de modèles *Biomodels*. Ce réseau est issu de (Schmierer et al., 2008), et représente les mécanismes de transduction de signal avec la molécule *Smad*, depuis la membrane cellulaire jusqu'au noyau de la cellule.

On ne considère ici qu'un sous-réseau de ce modèle, suffisant pour illustrer le cas de non confluence. Il est représenté dans la Fig. 5.10. L'espèce  $S4_c$ , qui représente la molécule de *Smad4* dans le cytoplasme, peut se lier soit avec une molécule de *Smad2* phosphorée ( $pS2_c$ ), formant le complexe  $S24_c$  (réaction  $r_5$ ), ou avec une molécule de  $G$  phosphorée ( $pG_c$ ), formant le complexe  $G4_c$  (réaction  $r_{22}$ ). Ces deux réactions sont réversibles ( $r_{5'}$  et  $r_{22'}$ ). Les mêmes types de transformations ont lieu dans le noyau (réactions  $r_6$ ,  $r_{6'}$ ,  $r_{23}$  et  $r_{23'}$ ). Le *Smad4* peut également se déplacer du cytoplasme vers le noyau, ou réciproquement ( $r_1$  et  $r_{1'}$ ). Enfin, le complexe de *Smad2* et *Smad4* peut se déplacer du cytoplasme vers le noyau ( $r_7$ ).

On suppose que  $\mathcal{J} = \{S4_c, S24_c, S4_n, S24_n\}$ . Le réseau est bien  $\mathcal{J}$ -linéaire. On peut donc éliminer ces quatre espèces. On obtient alors des réseaux avec la même structure, mais des cinétiques différentes, représentés dans la Fig. 5.11. Le premier réseau  $W_1$  est obtenu en supprimant d'abord  $S4_n$ , puis  $S24_n$ ,  $S24_c$ ,  $S4_c$  et les réactions dépendantes. Le deuxième  $WW_2$  est obtenu en supprimant  $S4_n$ , puis  $S24_n$ ,  $S4_c$ ,  $S24_c$  et les réactions dépendantes.

Fig 5.10: Sous-réseau du modèle *BIOMD0000000173*.

On note  $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}'_1, \mathbf{r}_5, \mathbf{r}'_5, \mathbf{r}_6, \mathbf{r}'_6, \mathbf{r}_7, \mathbf{r}_{22}, \mathbf{r}'_{22}, \mathbf{r}_{23}, \mathbf{r}'_{23})$  la base de réactions. Considérons le mode  $u = (1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0)$ . Il a deux décompositions en modes élémentaires,  $u_1 + u_2$  et  $u_3 + u_4$ , représentées dans la Fig. 5.12 (par simplicité, on omet les espèces non intermédiaires dans le graphe), avec :

$$\begin{aligned} u_1 &= (0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0) & u_2 &= (1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0) \\ u_3 &= (1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0) & u_4 &= (0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 0) \end{aligned}$$

### 5.3.5 Confluence de la simplification de systèmes d'équations

On s'intéresse à la simplification des systèmes d'équations dérivées de ces réseaux. On considère ici les règles (**VARIABLE STABLE**), (**VARIABLE CONSTANTE**), et (**SIMILARITÉ**) de la Fig. 4.13.

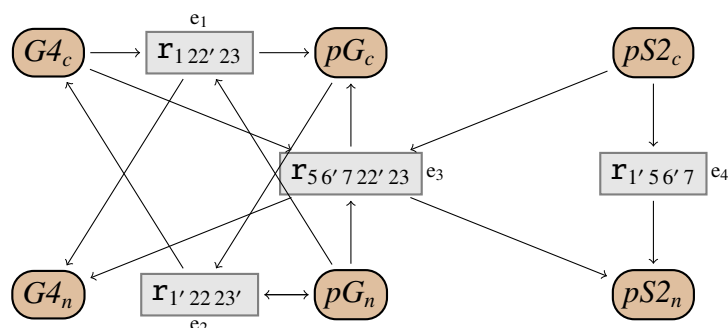
Cette simplification est trivialement uniformément confluente.

**Théorème 9.** La simplification de systèmes d'équations avec les trois règles ci-dessus est uniformément confluente dans  $(Systemes, \equiv)$ .

*Proof.* C'est direct, les règles ne faisant qu'appliquer des substitutions, qui commutent par similarité.  $\square$

Notons que comme il y a une relation de simulation entre la simplification de réseaux de flux et celle de réseaux de réactions, et une simulation entre cette simplification et celle des systèmes d'équations, il existe aussi une simulation entre la simplification des réseaux de flux et celle de systèmes d'équations.

**Proposition 18.** Soit  $\mathbf{r}$  une base de réactions, et  $W, W'$  deux réseaux de flux tel que  $W \Rightarrow W'$ . Alors  $E(W\mathbf{r}) \Rightarrow^* E(W'\mathbf{r})$ , en utilisant les trois règles ci-dessus.



	Réseau $W_1$	Réseau $W_2$
$e_1$	$\frac{2x_{G_{4c}}x_{pG_n}}{K_1}$	$\frac{x_{G_{4c}}x_{pG_n}(K_1 + K_3x_{pS_{2c}})}{K_1K_2}$
$e_2$	$\frac{2x_{G_{4n}}x_{pG_c}}{K_1}$	$\frac{2x_{G_{4n}}x_{pG_c}}{K_1}$
$e_3$	$\frac{x_{G_{4c}}x_{pG_n}x_{pS_{2c}}}{K_1}$	$\frac{x_{G_{4c}}x_{pG_n}x_{pS_{2c}}(x_{pG_c} + x_{pS_{2c}} + 1)K_3}{K_1K_2}$
$e_4$	$\frac{x_{pS_{2c}}(x_{G_{4c}} + x_{G_{4n}})}{K_1}$	$\frac{x_{pS_{2c}}(x_{G_{4c}}(x_{pS_{2c}} + x_{pG_c})K_3 + x_{G_{4n}}K_2)}{K_1K_2}$

Paramètres :

$$K_1 = 2x_{G_{4c}} + 2x_{pG_c} + 2x_{G_{4c}}x_{pG_c} + x_{pG_c}x_{pS_{2c}}$$

$$K_2 = x_{G_{4c}} + x_{pG_c} + x_{pS_{2c}} + x_{G_{4c}}x_{pG_c} + x_{pS_{2c}}x_{pG_c}$$

$$K_3 = 1 + x_{pG_c}$$

Fig 5.11: Réseaux simplifiés. Les deux réseaux ont la même structure (représentée graphiquement) mais des expressions cinétiques différentes (définies dans le tableau). Le réseau  $W_1$  est obtenu en éliminant dans l'ordre  $S_{4n}$ ,  $S_{24n}$ ,  $S_{24c}$ ,  $S_{4c}$  et les réactions dépendantes. Le réseau  $W_2$  est obtenu en éliminant dans l'ordre  $S_{4n}$ ,  $S_{24n}$ ,  $S_{4c}$ ,  $S_{24c}$ , et les réactions dépendantes.

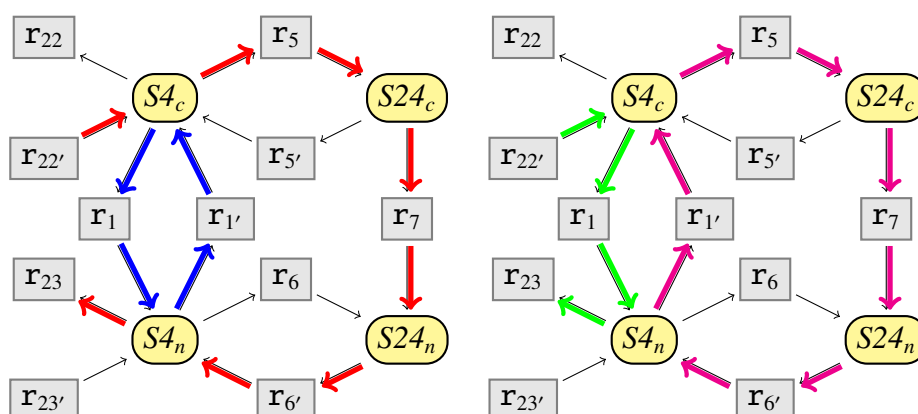


Fig 5.12: Deux décompositions du mode  $u$ , en les modes élémentaires  $u_1 + u_2$  et  $u_3 + u_4$ . Par simplicité, on a omis les espèces non intermédiaires.

*Proof.* C'est la conséquence de la Prop. 14, qui montre une simulation entre les réseaux de flux et les réseaux de réactions, et de la Prop. 13, qui montre une simulation entre les réseaux de réactions et les systèmes d'équations.  $\square$

Cela implique que, même lorsque la base de réactions n'est pas décomposable de façon unique, les réseaux complètement simplifiés ont le même système d'équations (en plus d'avoir la même structure).

**Théorème 10.** Soit  $W$  un réseau, et  $W_1, W_2$  deux réseaux tels que  $W \Rightarrow^* W_i$  et complètement simplifiés, c'est-à-dire  $\{W' \mid (W_i \Rightarrow W')\} = \emptyset$ . Alors  $W_1 \equiv^{struc} W_2$  et  $E(W_1\mathbf{r}) \equiv E(W_2\mathbf{r})$ .

*Proof.*  $W_1 \equiv^{struc} W_2$  est la conséquence directe du Théorème 7. De plus, d'après la Prop. 18, on a  $E(W\mathbf{r}) \Rightarrow^* E(W_i\mathbf{r})$ . Ces systèmes n'ont plus de variables pour les espèces intermédiaires, donc ne peuvent plus être simplifiés. Donc, d'après le Théorème 9,  $E(W_1\mathbf{r}) \equiv E(W_2\mathbf{r})$ .  $\square$

## 5.4 Discussion

On discute ici de la simplification confluente des réseaux de flux  $\mathcal{J}$ -linéaires, en particulier en les comparant aux méthodes de simplification proposées dans (Radulescu et al., 2008) et dans (Sáez et al., 2016).

Dans (Radulescu et al., 2008), Radulescu et al. reprend la méthode proposée par (Clarke, 1992) pour simplifier en trois étapes des réseaux, en éliminant des espèces intermédiaires. Premièrement, ils calculent le système simplifié d'équations, en supposant que les espèces intermédiaires sont stables. Ils utilisent pour cela des approximations, en supposant que certaines vitesses sont dominantes par rapport aux autres, ou en utilisant les équilibres tropicaux (Radulescu et al., 2015). Ils n'ont alors pas besoin de supposer que les cinétiques sont linéaires. Deuxièmement, ils calculent la structure du réseau simplifié en utilisant les modes élémentaires et la méthode de double description (Gagneur & Klamt, 2004). Enfin, ils cherchent à associer à chaque réaction simplifiée un taux cinétique, permettant d'obtenir le système d'équations calculé lors de la première étape. Cette étape n'est décrite formellement que dans le cas où les réactions simplifiées sont "stœchiométriquement indépendantes", c'est-à-dire qu'on ne peut pas par exemple avoir de réactions  $r_1, \dots, r_4$  telles que  $stoic_{r_1} + stoic_{r_2} = stoic_{r_3} + stoic_{r_4}$ . Lorsque cette condition est vérifiée, l'assignation des cinétiques est unique.

Leur méthode diffère donc de notre simplification sur plusieurs points. Premièrement, leur méthode est plus générale, elle ne nécessite pas de conditions particulières de linéarité ou de stœchiométrie. Elle supprime d'un seul coup un ensemble d'espèces intermédiaires, contrairement à notre simplification par règles, qui procède de façon récursive. Deuxièmement, leur simplification est une approximation, tandis que la notre est exacte. En ajoutant des conditions de linéarité similaires aux nôtres, il est probable qu'ils puissent également obtenir une simplification exacte, mais ils ne considèrent pas ce cas. Comme on a vu que nos règles de simplification calculaient aussi les modes élémentaires, la structure du réseau simplifié sera la même pour les deux méthodes de simplification. Les expressions cinétiques seront par contre différentes, puisqu'ils utilisent des approximations.

Concernant la confluence, leur méthode est trivialement structurellement confluente. Pour les cinétiques, elle n'est confluente que si les réactions simplifiées sont stœchiométriquement indépendantes. Cette condition est en fait équivalente à notre critère sur la décomposition de façon unique. Dans (Radulescu et al., 2012), ils proposent une étape supplémentaire dans le cas où les réactions ne seraient pas stœchiométriquement indépendantes. Cette étape supprime, de façon arbitraire, l'une des réactions non indépendantes, jusqu'à obtenir un ensemble indépendant (par exemple l'une des réactions  $r_i$  ci-dessus). L'assignation des cinétiques est alors unique. Mais comme cette étape d'élimination est arbitraire cette méthode



n'est pas confluente, et les réseaux simplifiés peuvent alors avoir différentes structures.

Dans (Sáez et al., 2016), les auteurs ont développé, en parallèle et indépendamment de nos travaux, une autre méthode de simplification de réseaux qui éliminent les espèces intermédiaires. Comme nous, leur simplification préserve exactement la sémantique déterministe lorsque les espèces intermédiaires sont stables. Leur méthode est basée sur la théorie des graphes et le calcul des arbres recouvrant dans le graphe des espèces intermédiaires.

En comparaison avec notre simplification, leur méthode cherche à supprimer en une seule étape un ensemble d'espèces intermédiaires. Les conditions sur les intermédiaires (linéarité, stœchiométrie, etc.) sont les mêmes pour les deux méthodes. Chaque réaction simplifiée avec leur méthode correspond à un chemin de taille minimal dans le graphe des intermédiaires, qui ne modifie pas la quantité des espèces intermédiaires. Il s'agit donc encore une fois de modes élémentaires, donc la structure du réseau simplifié sera la même que la notre ou celle de (Radulescu et al., 2008). Appliquée sur une unique espèce intermédiaire, leur méthode donne en fait exactement le même résultat que notre règle (INTERMÉDIAIRE - LINÉAIRE).

Leur simplification, appliquée sur un ensemble d'espèces intermédiaires, produit un unique réseau simplifié, et est donc directement confluente. Notons que le réseau produit par leur simplification semble toujours correspondre à l'un des réseaux produit par notre méthode. Il serait alors intéressant de mieux comprendre pourquoi ils obtiennent précisément ce réseau, et si celui-ci a des propriétés particulières par rapport aux autres réseaux qu'on peut obtenir avec notre simplification. Notons qu'ils proposent également d'appliquer leur méthode de façon récursive, sur une unique espèce intermédiaire à chaque étape. Dans ce cas, leur méthode est similaire à appliquer la règle (INTERMÉDIAIRE - LINÉAIRE), et n'est donc pas confluente (même pour la structure), puisque les réactions dépendantes ne sont pas supprimées. Le résultat de cette méthode n'est donc pas le même qu'avec leur méthode globale.

En résumé, il est donc possible d'obtenir une simplification exacte des espèces intermédiaires qui soit complètement confluente. Néanmoins, cela nécessite de nombreuses restrictions (linéarité, stœchiométrie, critère de décomposition unique), qui ne seront pas toujours vérifiées en pratique (voir l'exemple de la Sec. 5.3.4). Les conditions de linéarité et de stœchiométrie sont nécessaires pour que la simplification soit exacte, mais peuvent être relâchées si on souhaite juste une approximation (comme pour la simplification de (Radulescu et al., 2008)). Elles constituent néanmoins la plus grande restriction pour obtenir une simplification exacte et confluente. Le critère de décomposition unique, à l'inverse, n'est finalement que peu restrictif. En effet, en explorant l'ensemble de tous les réseaux

---

de la base de modèles [Biomodels](#) vérifiant les conditions de linéarité, et en considérant au maximum quatre espèces intermédiaires, nous n'avons trouvé qu'un unique modèle ne respectant pas le critère de décomposition unique, et donc pour lequel la simplification n'est pas confluente. Il s'agit du modèle présenté dans la [Sec. 5.3.4](#). Donc, dans la majorité des réseaux biologiques, le critère est vérifié. Notons également que si l'objectif est d'utiliser la simplification pour comparer des modèles, alors il n'est pas nécessaire de vérifier le critère de décomposition unique. On peut en effet simplifier les réseaux, puis comparer séparément leurs structures (on est structurellement confluente d'après le [Théoreme 7](#)) et leurs systèmes simplifiés d'équations (qui est aussi confluente d'après le [Théoreme 9](#)).

## 5.5 Conclusion

Dans ce chapitre, on a introduit un nouveau type de réseau, les réseaux de flux cinétiques, et une nouvelle règle de simplification supprimant les flux dépendants. On a montré que si on néglige les cinétiques, la simplification des espèces intermédiaires est confluente et correspond au calcul des modes élémentaires. Avec les cinétiques, il est nécessaire d'ajouter des conditions sur la linéarité des cinétiques et la stœchiométrie des espèces intermédiaires. La simplification supprime alors l'ensemble des espèces intermédiaires et est structurellement confluente. On donne également un critère sur le réseau initial garantissant que la simplification soit complètement confluente, y compris pour les cinétiques. On prouve également que la simplification de systèmes d'équations est confluente.

Cette méthode de simplification confluente pourrait alors être utilisée de façon systématique sur les modèles de [Biomodels](#), pour les comparer. Il serait également intéressant d'étudier plus en détails les liens entre notre simplification et celle proposée dans ([Sáez et al., 2016](#)), afin de comprendre pourquoi leur méthode donne directement un unique réseau.

## CHAPITRE 6

---

### Discussion

---

Dans ce chapitre, on s'intéresse à l'uniformisation de nos règles de simplification pour différentes sémantiques. Dans un premier temps, on compare les sémantiques et les règles des chapitres 3 et 4. On montre que même si on retrouve dans les deux cas les mêmes types de simplification, il n'est pas toujours possible d'obtenir des règles correctes pour les deux sémantiques. En particulier, on ne peut pas obtenir de règle unifiée pour l'élimination des espèces intermédiaires. On discute ensuite de la généralisation de ces règles à une sémantique stochastique, et en particulier des difficultés supplémentaires qu'elle apporte.

## 6.1 Des règles de simplification préservant les deux sémantiques ?

Dans les chapitres 3 et 4, on a présenté deux méthodes de simplification, préservant respectivement la sémantique non déterministe des attracteurs et la sémantique déterministe basée sur les systèmes d'équations différentielles ordinaires. Les règles des deux simplifications sont basées sur les mêmes principes : élimination d'espèces intermédiaires, suppression de réactions combinatoires d'autres réactions, utilisation des symétries, etc. Il est donc naturel de s'interroger sur l'existence d'une méthode de simplification unique qui préserverait les deux sémantiques. On discute ici (informellement) d'une telle simplification, et plus généralement des différences entre les différentes sémantiques.

Un réseau, dans ce cas, est alors composé d'un ensemble de réactions cinétiques, d'une contrainte (pour la sémantique déterministe), et d'une solution chimique initiale (pour la sémantique non déterministe). Les réactions ne peuvent pas être normalisées, car cela modifierait la sémantique non déterministe (par exemple, une réaction  $A + B \rightarrow C + B$ ;  $k_1 x_A x_B$  n'est pas applicable en sémantique non déterministe dans la solution  $A$ , tandis que la réaction normalisée  $A \rightarrow C$ ;  $k_1 x_A x_B$  devient applicable). En fait, la normalisation ne sera possible qu'avec la règle (**ACTIVATEUR PRÉSENT**), lorsque l'activateur est toujours présent dans tous les réseaux accessibles.

Certaines règles de simplification peuvent alors facilement être uniformisées pour préserver les deux sémantiques. C'est notamment le cas par exemple des règles (**RÉACTION CYCLIQUE**), (**CONTEXTE**), ou encore (**ESPÈCE INUTILE**).

Mais les deux sémantiques présentent néanmoins d'importantes différences, en terme de traitement des quantités de matières, du temps, des conditions initiales, etc. Dans la suite, on étudie ces différences et leurs conséquences sur les règles de simplification. En particulier, certaines règles ne seront valables que pour certaines sémantiques.

Une première distinction majeure entre les deux sémantiques est le traitement des *quantités de matières*. Dans la sémantique non déterministe, on considère des populations de molécules, tandis que la sémantique déterministe utilise des concentrations. En particulier, l'approche avec des concentrations suppose que le nombre de molécules est grand, et on ne considère jamais le cas où une espèce est présente, mais en quantité insuffisante pour appliquer une réaction. Par exemple, considérons une réaction  $2A \rightarrow 2B$ ;  $k x_A$ . Avec une sémantique déterministe, soit le réseau ne contient pas de  $A$ , et la réaction n'est pas applicable, soit le réseau contient du  $A$ , et elle est applicable. En non déterministe, il y a une troisième pos-

sibilité : il y a une unique molécule de  $A$ , et la réaction n'est donc pas applicable.

Cette différence a deux conséquences pour les règles de simplification. Premièrement, les règles qui simplifient la stœchiométrie des réactions ne sont valables que pour la sémantique déterministe. C'est par exemple le cas de la règle (**FACTORISATION**), qui remplace une réaction  $2A \rightarrow 2B; k_{x_A}$  par une réaction  $A \rightarrow B; 2k_{x_A}$ . Deuxièmement, certaines règles pour le cas non déterministe ont besoin de conditions supplémentaires pour que le réseau ne soit pas "bloqué", comme on l'a vu par exemple dans la règle (**INTERMÉDIAIRE - PRODUIT**).

Une autre différence, évidente, est la présence d'expressions *cinétiques* pour la sémantique déterministe. Comme la sémantique des attracteurs n'utilisent pas ces cinétiques, les règles qui ne font que les modifier, comme par exemple (**MODIFICATEUR CONSTANT**) ou (**SIMILARITÉ**), seront directement valables pour les deux sémantiques. À l'inverse, certaines règles pour le cas non déterministe nécessitent d'ajouter des conditions sur les cinétiques pour être correctes. Par exemple, considérons la règle (**PRODUCTIONS GROUPÉES**), qui peut par exemple fusionner la réaction  $\emptyset \rightarrow A$  avec la réaction  $\emptyset \rightarrow B$ , formant la réaction  $\emptyset \rightarrow A + B$ . Pour être valable dans la sémantique déterministe, il faut alors ajouter comme condition que les deux réactions initiales aient la même expression cinétique.

Enfin, une autre différence majeure concerne les *quantités initiales*. Pour la sémantique non déterministe, on considère un réseau avec une solution initiale *fixée*. Cette condition est nécessaire pour pouvoir appliquer la simplification. En effet, 15 des 20 règles du cas non déterministe ont besoin de conditions sur la solution initiale, ou bien doivent la modifier. À l'inverse, dans la sémantique déterministe, on est plus général, et les simplifications sont valables *quelque soit* la concentration initiale (modulo éventuellement les contraintes). Notons qu'il faudrait donc ajouter une condition de cohérence aux réseaux généralisés pour les deux sémantiques, afin d'éviter d'avoir par exemple un réseau sans molécule de  $A$  dans sa solution initiale et en même temps une contrainte  $x_A \neq 0$ .

En particulier, cette différence rend impossible dans la majorité des cas d'avoir une règle unifiée d'*élimination des espèces intermédiaires*. En effet, pour la sémantique des attracteurs, la plupart des règles d'élimination d'une espèce imposent qu'elle ne soit pas présente dans la solution initiale. Pour la sémantique déterministe, on suppose par contre que l'espèce intermédiaire est dans un état d'équilibre, et avec une concentration non nulle pour que la simplification soit intéressante. Le seul cas où on peut unifier ces règles est celui où il y a une unique réaction de consommation, de forme  $A \rightarrow s; e'$ , avec  $A$  l'espèce intermédiaire (voir la règle (**INTERMÉDIAIRE - SOLUTION**)). Dans ce cas, toute réaction de production  $s_1 \rightarrow s_2 + aA; e$  sera remplacée par la réaction  $s_1 \rightarrow s_2 + as; e$ , avec exactement la même expression cinétique.

## 6.2 Simplification en stochastique

On s'intéresse (brièvement) à l'extension des règles de simplification à une sémantique stochastique, notamment en étudiant les méthodes de simplification existantes dans la littérature.

La *sémantique stochastique* décrit la dynamique probabiliste des réseaux. À tout temps  $t \in \mathbb{R}_{\geq 0}$ , elle associe une distribution de probabilité entre les différentes solutions dans lesquelles peut être le modèle. Elle peut alors être décrite par une *chaîne de Markov à temps continu* (Norris, 1998). L'*équation maîtresse chimique* (D. T. Gillespie, 1977), un système d'équations différentielles sur la probabilité d'être dans une certaine solution au temps  $t$ , permet également de décrire l'évolution stochastique du système. De façon similaire au cas déterministe, une *solution* du réseau en sémantique stochastique est alors une solution de son équation maîtresse. Deux réseaux sont alors équivalents s'ils ont les mêmes solutions observables.

La plupart des simplifications existantes dans la littérature pour la sémantique stochastique sont des approximations. Par exemple, dans (Batmanov, Kuttler, Lemaire, Lhoussaine, & Versari, 2012), les auteurs utilisent des méthodes de clôture de moments (C. S. Gillespie, 2009) pour approximer l'équation maîtresse, puis se servent de symétrie pour simplifier le réseau.

Une version stochastique de l'hypothèse de quasi-équilibre est proposée dans (C. V. Rao & Arkin, 2003). Soient  $X_1, \dots, X_k$  les espèces intermédiaires, et  $A_1, \dots, A_l$  les autres espèces. On note

$$P((n_1, \dots, n_k) | (a_1, \dots, a_l), t),$$

la probabilité conditionnelle d'être au temps  $t$  dans un état contenant  $n_1$  molécules de  $X_1, \dots$ , et  $n_k$  molécules de  $X_k$ , sachant qu'il y a  $a_1$  molécules de  $A_1, \dots$ , et  $a_l$  molécules de  $A_l$ . Alors les espèces intermédiaires sont à l'équilibre stochastique si ces probabilités ne dépendent pas du temps, c'est-à-dire si pour tout  $n_1, \dots, n_k$ , pour tout  $a_1, \dots, a_l$ , on a :

$$\frac{dP((n_1, \dots, n_k) | (a_1, \dots, a_l), t)}{dt} = 0.$$

Dans (C. V. Rao & Arkin, 2003), l'hypothèse faite est qu'on est proche d'un tel équilibre. Ils considèrent alors l'exemple classique de Michaelis-Menten, et montrent que si les espèces intermédiaires  $\mathcal{J} = \{E, C\}$  sont proche de l'équilibre stochastique, alors le comportement du réseau simplifié est une approximation de celui du réseau initial, avec une sémantique stochastique. Leur simplification

n'est néanmoins pas étendue à des cas plus généraux d'espèces intermédiaires.

Considérons maintenant les règles de simplification proposées dans cette thèse pour la sémantique des attracteurs et la sémantique déterministe.

Certaines règles s'étendent trivialement à la sémantique stochastique. Pour la sémantique des attracteurs, c'est par exemple le cas de (**RÉACTION CYCLIQUE**) ou d'(**ACTIVATEUR PRÉSENT**). Les règles (**MODIFICATEUR CONSTANT**) ou (**SIMILARITÉ**) de simplification avec une sémantique déterministe sont également directement valables pour le cas stochastique. En adaptant les conditions à la sémantique stochastique, il semble également possible d'étendre les règles (**ESPÈCE INUTILE**) et (**RÉACTIONS INAPPLICABLES**).

À l'inverse, certaines règles sont directement fausses en stochastique. C'est par exemple le cas de la règle (**FACTORISATION**), ou bien de la règle (**PRODUCTIONS GROUPÉES**).

Considérons plus en détails le cas de l'élimination d'espèces intermédiaires. On reprend la simplification proposée dans (C. V. Rao & Arkin, 2003), mais en considérant un équilibre stochastique exact. La simplification de Michaelis-Menten conserve alors exactement la sémantique stochastique (lorsqu'il n'y a pas de contexte, comme expliqué ci-dessous). Il semble alors possible de généraliser cette simplification à d'autres cas d'espèces intermédiaires, comme on l'a fait pour le cas déterministe avec les règles de la Fig. 4.3.

La question du contexte apporte des difficultés supplémentaires dans la sémantique stochastique par rapport aux deux autres sémantiques. Reprenons l'exemple de la simplification de Michaelis-Menten. En étudiant les calculs proposés dans (C. V. Rao & Arkin, 2003), il apparaît qu'ils utilisent en fait une autre hypothèse, celle que la probabilité d'avoir  $a$  molécules du substrat  $S$  ne dépend pas du nombre de molécules du produit  $P$ , c'est-à-dire qu'il n'y a pas de boucle de rétro-action. Plus formellement, pour tout  $a, b, t$ , il faut que :

$$P(aS \mid bP, t) = P(aS, t).$$

Cette hypothèse est trivialement vérifiée dans leur cas, puisqu'ils étudient le système Michaelis-Menten isolé, c'est-à-dire sans aucun contexte. Elle ne peut en revanche jamais être vérifiée si on considère une équivalence dans tout contexte.

Donc, afin d'avoir une sémantique stochastique contextuelle avec laquelle on puisse faire des simplifications d'espèces intermédiaires, il est nécessaire de raffiner la notion de contexte. Une solution envisageable serait alors de partitionner les espèces en quatre ensembles différents : les espèces complètement libres, les espèces complètement internes, les espèces en entrée et les espèces en sortie. Le contexte serait lui aussi divisé en trois : un contexte global, qui ne peut modifier



que les espèces complètement libres; un contexte d'entrée, qui ne peut modifier que les espèces en entrée; et un contexte de sortie, qui ne peut modifier que les espèces en sortie. Pour Michaelis-Menten par exemple,  $S$  serait une espèce en entrée,  $E$  et  $C$  des espèces complètement internes, et  $P$  une espèce en sortie.

## 6.3 Conclusion

Dans ce chapitre, on a étudié une extension des règles de simplification, afin de pouvoir soit préserver en même temps la sémantique des attracteurs et la sémantique déterministe, soit pouvoir préserver la sémantique stochastique. On a vu que certaines règles pouvaient trivialement être étendues, tandis que d'autres sont vraiment spécifiques à une sémantique donnée. En particulier, bien qu'il semble possible d'éliminer les espèces intermédiaires dans les trois sémantiques, les conditions varient beaucoup d'un cas à l'autre, et il est donc difficile d'avoir une méthode de simplification unifiée.



## CHAPITRE 7

---

### Conclusion et perspectives

---

L'objectif de cette thèse était de pouvoir *simplifier* des systèmes biologiques, tout en *préservant leurs sémantiques*.

Nous avons tout d'abord proposé un ensemble de *règles de simplification* pour le cas *non déterministe*. Ces règles *structurelles* permettent de simplifier des réseaux de réactions sans cinétiques, notamment en supprimant des espèces intermédiaires. En particulier, nous avons montré qu'en se restreignant à un sous ensemble de 6 règles, la simplification était *confluente*, et suffisante pour réduire efficacement le réseau Tet-On. Nous avons défini une nouvelle sémantique *contextuelle*, la *sémantique des attracteurs*, basée sur la capacité d'un réseau de diverger, ou de converger vers des attracteurs. Deux réseaux équivalents auront alors, dans tout contexte, les mêmes *comportements finaux réguliers*. Nous avons prouvé que la simplification préserve cette sémantique, et donc par exemple que le réseau initial de Tet-On était équivalent à son réseau simplifié. Nous avons également montré sur le système MAPK que nous pouvions utiliser la simplification pour comparer différents modèles d'un même système biologique.

Pour la *sémantique déterministe*, nous avons proposé de nouvelles *règles de simplification* pour les réseaux avec cinétiques. Ces règles sont *structurelles*, c'est-à-dire que la simplification d'un réseau est encore un réseau, et sont *symboliques*, c'est-à-dire ne nécessitent pas de connaître la valeur des différents paramètres cinétiques. Elles calculent en même temps la structure des réactions simplifiées et leurs nouvelles expressions cinétiques. Elles préservent, dans tout *contexte*, le comportement déterministe des réseaux. Par exemple, des règles permettent d'éliminer des réactions intermédiaires, lorsqu'elles sont à l'équilibre. Ces règles

sont suffisantes pour pouvoir simplifier le système Tet-On. Nous avons également montré comment adapter ces règles pour les appliquer à des systèmes d'équations plutôt qu'à des réseaux.

Enfin, nous avons étudié la *confluence* de l'*élimination des espèces intermédiaires* pour la sémantique déterministe. Nous avons montré que, en négligeant les cinétiques, l'élimination des espèces intermédiaires et des réactions dépendantes était une simplification confluente, et qu'elle permettait de calculer les *modes élémentaires* d'un réseau. Avec les cinétiques, nous avons montré qu'il fallait ajouter des conditions supplémentaires. Lorsque les espèces intermédiaires sont *linéaires* et que leurs stœchiométries vérifient certaines conditions, il est alors possible de toutes les supprimer, et les réseaux simplifiés ont alors tous la *même structure*. Si en plus le réseau initial vérifie un critère de décomposition unique, ces réseaux simplifiés ont également les *même cinétiques*, et donc la simplification est complètement *confluente*.

Les travaux présentés dans cette thèse peuvent être continués et étendus de différentes façons. Nous proposons ici quelques pistes de réflexion.

Premièrement, il serait intéressant d'*appliquer* les méthodes de simplification présentées dans cette thèse sur d'autres réseaux biologiques, par exemple ceux de la base de modèle [Biomodels](#). Pour chaque modèle, il faut alors définir quelles sont les espèces internes, quelles est la fonction d'observation, quelles espèces peuvent être à l'équilibre, etc. La simplification peut alors être utilisée pour *comparer* ces modèles, comme nous l'avons fait pour MAPK. En étudiant les réseaux simplifiés, il sera alors peut-être possible de trouver de nouvelles règles de simplification, et donc d'obtenir une simplification plus *complète*. Il est également possible d'essayer d'élargir les propriétés de *confluence* à plus de règles de simplification.

Une autre possibilité consiste à étendre nos règles à d'*autres sémantiques* et d'*autres langages* de modélisation. Comme discuté dans le Chapitre 6, il semble possible d'avoir le même type de simplification pour une sémantique *stochastique*. Il faut néanmoins définir formellement les règles, et la nouvelle notion de contexte pour pouvoir prendre en compte les boucles de rétro-action. Il est également possible d'essayer de relâcher les conditions d'équilibre pour la sémantique stochastique. Dans ce cas là, le comportement du réseau simplifié serait une *approximation* de celui du réseau initial. Il faut alors définir formellement une notion de correction, permettant d'estimer et de *borner l'erreur* introduite par la simplification. Nos règles ont été développées pour des réseaux de réactions. Il est néanmoins possible de modéliser les systèmes biologiques à l'aide d'autres langages de modélisation. Il est alors naturel d'essayer d'étendre nos règles à ces autres langages, par exemple pour les langages à base de règles comme *Kappa* ([Danos & Laneve, 2004](#)) ou *React(C)* ([John et al., 2011](#)).

---

## Liste des figures

---

1.1	Exemple de réactions chimiques. La réaction $r_1 = A \rightarrow B$ transforme une molécule de l'espèce $A$ en une molécule de l'espèce $B$ . La réaction $r_2 = A + C \rightarrow D$ transforme une molécule de $A$ et une molécule de $C$ en une molécule de $D$ . . . . .	13
1.2	Exemple de réactions chimiques avec cinétiques. La vitesse $k_1 x_A$ de la réaction $r_1 = A \rightarrow B$ ; $k_1 x_A$ dépend d'un paramètre $k_1$ et de la quantité $x_A$ de l'espèce $A$ . La vitesse de la réaction $r_2 = A + C \rightarrow D$ ; $k_2 x_A x_C$ est $k_2 x_A x_C$ . . . . .	14
1.3	Réseau de réactions cinétiques du système Tet-On. . . . .	15
1.4	Graphe d'accessibilité pour les réactions de la Fig. 1.1, avec initialement deux molécules de $A$ et une molécule de $C$ . Les cercles pointillés représentent les états terminaux. . . . .	16
1.5	Exemple de système d'équations différentielles et de l'évolution des concentrations des réactions de la Fig. 1.1, pour des valeurs de paramètres et des concentrations initiales fixées. La concentration de $A$ est en vert, celle de $B$ en rouge, celle de $C$ en bleu et celle de $D$ en magenta. . . . .	17
1.6	Exemple d'une simulation stochastique des réactions de la Fig. 1.2, avec initialement 50 molécules de $A$ et 25 molécules de $C$ . Le nombre de molécules de $A$ est en vert, le nombre de $B$ en rouge, celui de $C$ en bleu et celui de $D$ en magenta. . . . .	18
1.7	Réseau initial (gauche) et simplifié (droite) pour Michaelis-Menten. . . . .	20
1.8	Exemple de simplification du réseau de gauche pour obtenir le réseau de droite. Les espèces $X$ , $Y_1$ et $Y_2$ sont des espèces intermédiaires, que l'on souhaite supprimer. . . . .	23

1.9	Réseau initial du système Tet-On. . . . .	26
1.10	Réseau simplifié du système Tet-On. . . . .	26
1.11	Exemple de simplification d'une espèce intermédiaire $X$ . La réaction $r_{1^3 3^2}$ est la combinaison de 3 fois la réaction $r_1$ et de 2 fois la réaction $r_3$ . Elle consomme trois molécules de $A$ , produit 2 molécule de $C$ , et ne modifie pas la quantité de $X$ . . . . .	27
1.12	Réseau partiellement simplifié du système Tet-On, après avoir supprimé l'espèce intermédiaire $GFP$ . . . . .	27
1.13	À gauche : graphe des réactions. À droite : graphe des état accessibles. Le réseau peut soit converger vers $\{B, C\}$ , soit diverger. . . . .	28
1.14	Réseau cinétique initial du système Tet-On. . . . .	30
1.15	Réseau cinétique simplifié du système Tet-On. . . . .	30
1.16	Modes élémentaires du système Tet-On (sans cinétique). . . . .	31
1.17	Réseau simplifié (sans cinétique) du système Tet-On. . . . .	31
1.18	Réseau de signalisation pour l'espèce $Smad$ , avec un mode décomposable en modes élémentaires de deux façons différentes. . . . .	32
2.1	Représentation graphique des réactions $r_1 = 2A + B \rightarrow C + B$ et $r_2 = A \rightarrow B$ . . . . .	39
2.2	Représentation graphique du réseau $N$ seul (à gauche) ou dans un contexte (à droite). . . . .	41
2.3	Graphe des solutions accessibles du réseau $N$ . . . . .	43
2.4	Graphe des solutions accessibles du réseau $\{\emptyset \rightarrow A\} \mid \emptyset$ . . . . .	43
2.5	Graphe du réseau de Michaelis-Menten. . . . .	44
2.6	Graphe des solutions accessibles du réseau de Michaelis-Menten. . . . .	44
2.7	Graphe du système Tet-On. . . . .	45
2.8	Réactions du réseau Tet-On. . . . .	45
2.9	Réactions du système MAPK. . . . .	46
2.10	Graphe du système MAPK. . . . .	47
2.11	Expressions, avec $x \in Vars, k \in Param, a \in \mathbb{R}$ . . . . .	51
2.12	Contraintes, avec $e, e' \in Expr$ . . . . .	53
2.13	Représentation graphique du réseau cinétique $M$ . . . . .	54
2.14	Système d'équations pour le réseau $M$ . . . . .	55
2.15	Syntaxe et interprétation d'un système d'équations. . . . .	55
2.16	Ces deux systèmes sont équivalents, si on observe uniquement la somme des concentration de $A, A_1$ et $A_2$ , et la concentration de $B$ . . . . .	56
2.17	Dérivation du système d'équations d'un réseau cinétique. . . . .	57
2.18	Ces deux réseaux sont équivalents, si on observe uniquement la somme des concentration de $A, A_1$ et $A_2$ , et la concentration de $B$ . . . . .	57
2.19	Graphe et système du réseau cinétique de Michaelis-Menten. . . . .	58
2.20	Graphe du réseau cinétique Tet-On. . . . .	59

2.21	Schéma d'une règle de simplification. . . . .	60
2.22	Diagrammes de la simulation $f$ . . . . .	61
2.23	Diagrammes de confluence pour la simplification $\rightarrow$ . . . . .	62
2.24	Exemples de confluence locale sans confluence globale. . . . .	62
3.1	Graphes des solutions accessibles des réseaux $N_1$ à $N_{12}$ . . . . .	69
3.2	Ensemble confluent de règles de simplification pour la congruence des attracteurs, avec la fonction d'observation $\Omega$ et les espèces internes $\mathcal{I}$ . . . . .	76
3.3	Graphe du réseau simplifié $N'_{\text{Tet-On}}$ . . . . .	92
3.4	Graphe du réseau intermédiaire lors de la simplification de Tet-On. . . . .	93
3.5	Règles de simplification d'espèces intermédiaires pour la congruence des attracteurs, avec la fonction d'observation $\Omega$ et les espèces internes $\mathcal{I}$ . . . . .	95
3.6	Simplification de cycles avec les règles (RÉDUCTION INVERSIBLE) et (RÉACTION CYCLIQUE). . . . .	97
3.7	Graphes des réseaux intermédiaires lors de la simplification de Michaelis-Menten. . . . .	98
3.8	Graphe du réseau simplifié de Michaelis-Menten. . . . .	99
3.9	Règles de décomposition pour la congruence des attracteurs, avec la fonction d'observation $\Omega$ et les espèces internes $\mathcal{I}$ . . . . .	102
3.10	Règles de symétrie pour la congruence des attracteurs, avec la fonction d'observation $\Omega$ et les espèces internes $\mathcal{I}$ . . . . .	105
3.11	Autres règles de simplification pour la congruence des attracteurs, avec la fonction d'observation $\Omega$ et les espèces internes $\mathcal{I}$ . . . . .	107
3.12	Classification des réseaux du système MAPK, avec $n$ le nombre d'espèce et $m$ celui de réactions. . . . .	123
3.13	Graphe du réseau simplifié de MAPK. . . . .	123
3.14	Exemple de réseaux attracteur-équivalents et non bisimilaires. . . . .	129
3.15	Réseau de réactions $N$ . Le réseau $N_1$ est obtenu en éliminant la réaction $r_2$ , et le réseau $N_2$ en fusionnant les espèces $B_1$ et $B_2$ . . . . .	130
3.16	Simplification du réseau $N$ en $N'$ avec la règle (INTERMÉDIAIRE - RÉACTANT). . . . .	131
4.1	Axiomes de similarité pour les expressions et les contraintes. . . . .	139
4.2	Axiomes de similarité pour les systèmes d'équations et les réseaux cinétiques. . . . .	140
4.3	Règles de simplification d'espèce intermédiaire en déterministe, avec $\mathcal{I}$ l'ensemble des espèces internes. . . . .	142
4.4	Graphes des réseaux cinétiques intermédiaire lors de la simplification de Michaelis-Menten. . . . .	145



4.5	Règle de simplification de combinaison en déterministe, avec $\Omega$ la fonction d'observation et $\mathcal{I}$ les espèces internes. . . . .	147
4.6	Règles de simplification d'espèces symétriques en déterministe, avec $\Omega$ la fonction d'observation et $\mathcal{I}$ les espèces internes. . . . .	148
4.7	Autres règles de simplification en déterministe, avec $\Omega$ la fonction d'observation et $\mathcal{I}$ l'ensemble des espèces internes. . . . .	151
4.8	Graphe du réseau cinétique simplifié de Michaelis-Menten, avec $V = k_3(x_C^0 + x_E^0)$ et $K = (k_2 + k_3)/k_1$ . . . . .	152
4.9	Graphe du réseau cinétique Tet-On. . . . .	159
4.10	Graphe d'un réseau cinétique partiellement simplifié de Tet-On. . . . .	160
4.11	Graphe d'un réseau cinétique partiellement simplifié de Tet-On. . . . .	160
4.12	Graphe du réseau cinétique simplifié de Tet-On. . . . .	161
4.13	Règles de simplification de systèmes d'équations, avec $\Omega$ la fonction d'observation. . . . .	163
5.1	Élimination des espèces intermédiaires $X$ et $Y$ dans les réactions $R$ , dans différents ordres, illustrant le besoin d'élimination des réactions dépendantes. . . . .	174
5.2	Règles de simplification des réseaux de flux sans cinétique, avec $\mathcal{I}$ l'ensemble des espèces internes. . . . .	176
5.3	Élimination des espèces intermédiaires $X$ , $Y$ et $Z$ dans le réseau $U$ , dans deux ordres différents, illustrant le besoin de la règle (FACTORISATION). . . . .	177
5.4	Matrice de stœchiométrie et graphe des réactions $R$ , avec en rouge le mode élémentaire $x_1 = (1, 0, 0, 2)$ et en bleu le mode élémentaire $x_2 = (0, 1, 1, 0)$ . . . . .	178
5.5	Axiomes de similarité pour les réseaux de flux cinétiques. . . . .	186
5.6	Réseau $W$ (en haut et à gauche) et ses simplifications : en haut à droite le réseau $W_{XYZ}$ obtenu en supprimant dans l'ordre les espèces intermédiaires $X$ , $Y$ et $Z$ ; en bas à gauche, le réseau $W_{XZY}$ obtenu en supprimant $X$ puis $Z$ puis $Y$ ; en supprimant le flux dépendant $u_{123456}$ du réseau $W_{XYZ}$ , on peut obtenir soit le réseau $W_{XYZd_1} = W_{XZY}$ , soit le réseau $W_{XYZd_2}$ en bas à droite. Ces deux réseaux ont la même structure, mais pas les mêmes expressions cinétiques. Le nouveau paramètre $K$ dans les expressions cinétiques est défini par $K = k_2k_3 + k_3k_4 + k_4k_5$ . . . . .	187
5.7	Règles de simplification de réseaux de flux cinétiques. . . . .	189
5.8	Le réseau initial $W_0$ (gauche) et le réseau simplifié $W$ (droite) après avoir enlevé $Y$ . . . . .	193
5.9	Deux décompositions du mode $u = (1, 1, 1, 1, 1, 1)$ dans le réseau $W$ . . . . .	204

- 
- 5.10 Sous-réseau du modèle *BIOMD0000000173*. . . . . 212
- 5.11 Réseaux simplifiés. Les deux réseaux ont la même structure (représentée graphiquement) mais des expressions cinétiques différentes (définies dans le tableau). Le réseau  $W_1$  est obtenu en éliminant dans l'ordre  $S4_n, S24_n, S24_c, S4_c$  et les réactions dépendantes. Le réseau  $W_2$  est obtenu en éliminant dans l'ordre  $S4_n, S24_n, S4_c, S24_c$ , et les réactions dépendantes. . . . . 213
- 5.12 Deux décompositions du mode  $u$ , en les modes élémentaires  $u_1 + u_2$  et  $u_3 + u_4$ . Par simplicité, on a omis les espèces non intermédiaires. 214

## Bibliographie

- Alla, H., & David, R. (1998). A modelling and analysis tool for discrete events systems: continuous petri net. *Performance Evaluation*, 33(3), 175–199.
- Alur, R., Belta, C., Ivančić, F., Kumar, V., Mintz, M., Pappas, G. J., ... Schug, J. (2001). Hybrid modeling and simulation of biomolecular networks. In *International workshop on hybrid systems: Computation and control* (pp. 19–32).
- Anderson, J., & Papachristodoulou, A. (2009). On validation and invalidation of biological models. *BMC bioinformatics*, 10(1), 1.
- Andrianantoandro, E., Basu, S., Karig, D. K., & Weiss, R. (2006). Synthetic biology: new engineering rules for an emerging discipline. *Molecular systems biology*, 2(1).
- Araki, T., & Kasami, T. (1977). Decidable problems on the strong connectivity of petri net reachability sets. *Theoretical Computer Science*, 4(1), 99–119.
- Ashyraliyev, M., Fomekong-Nanfack, Y., Kaandorp, J. A., & Blom, J. G. (2009). Systems biology: parameter estimation for biochemical models. *Febs Journal*, 276(4), 886–902.
- Autant, C., & Schnoebelen, P. (1992). Place bisimulations in petri nets. In *International conference on application and theory of petri nets* (pp. 45–61).
- Batmanov, K., Kuttler, C., Lemaire, F., Lhoussaine, C., & Versari, C. (2012). Symmetry-based model reduction for approximate stochastic analysis. In *Computational methods in systems biology* (pp. 49–68).
- Bause, F., & Kritzinger, P. S. (2002). *Stochastic petri nets*. Vieweg Wiesbaden.
- Berthelot, G. (1985). Checking properties of nets using transformations. In *European workshop on applications and theory in petri nets* (pp. 19–40).
- Berthelot, G., & Roucairol, G. (1976). Reduction of petri-nets. In *International symposium on mathematical foundations of computer science* (pp. 202–209).
- Berthelot, G., Roucairol, G., & Valk, R. (1980). Reductions of nets and parallel programs. In *Net theory and applications* (pp. 277–290). Springer.
- Bilke, S., & Sjunnesson, F. (2001). Stability of the kauffman model. *Physical Review E*, 65(1), 016129.
- Biomodels : base de données de réseaux biologiques*. (2016). (<http://biomodels.net/>)
- Blinov, M. L., Faeder, J. R., Goldstein, B., & Hlavacek, W. S. (2004). Bionetgen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics*, 20(17), 3289–3291.
- Bodenstein, M. (1913). Eine theorie der photochemischen reaktionsgeschwindigkeiten. *Z. phys. Chem*, 85(329), 0022–3654.

- Briggs, G. E., & Haldane, J. B. S. (1925). A note on the kinetics of enzyme action. *Biochemical journal*, 19(2), 338.
- Burgard, A. P., Pharkya, P., & Maranas, C. D. (2003). Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnology and bioengineering*, 84(6), 647–657.
- Butte, A. J., & Kohane, I. S. (2000). Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. In *Pac symp biocomput* (Vol. 5, pp. 418–429).
- Cardelli, L. (2014). Morphisms of reaction networks that couple structure to function. *BMC systems biology*, 8(1), 84.
- Cardelli, L., Tribastone, M., Tschaikowski, M., & Vandin, A. (2015). Forward and backward bisimulations for chemical reaction networks. *CONCUR*, 226–239.
- Cardelli, L., Tribastone, M., Tschaikowski, M., & Vandin, A. (2016). Symbolic computation of differential equivalences. *POPL*.
- Cha, S. (1968). A simple method for derivation of rate equations for enzyme-catalyzed reactions under the rapid equilibrium assumption or combined assumptions of equilibrium and steady state. *Journal of biological chemistry*, 243(4), 820–825.
- Chaouiya, C. (2007). Petri net modelling of biological networks. *Briefings in bioinformatics*, 8(4), 210–219.
- Chou, K.-C. (1981). Two new schematic rules for rate laws of enzyme-catalysed reactions. *Journal of theoretical biology*, 89(4), 581–592.
- Ciocchetta, F., & Hillston, J. (2008). Process algebras in systems biology. In *International school on formal methods for the design of computer, communication and software systems* (pp. 265–312).
- Clarke, B. L. (1980). *Stability of complex reaction networks*. Wiley Online Library.
- Clarke, B. L. (1992). General method for simplifying chemical networks while preserving overall stoichiometry in reduced mechanisms. *The Journal of chemical physics*, 97(6), 4066–4071.
- Colom, J., Martinez, J., & Silva, M. (1986). Packages for validating discrete production systems modelled with petri nets. In *Proceedings of the imacs-ifac symposium*.
- Cornish-Bowden, A. (1977). An automatic method for deriving steady-state rate equations. *Biochemical Journal*, 165(1), 55–59.
- Cornish-Bowden, A. (2012). Fundamentals of enzyme kinetics.
- Coutte, F., Niehren, J., Dhali, D., John, M., Versari, C., & Jacques, P. (2015). Modeling leucine's metabolic pathway and knockout prediction improving the production of surfactin, a biosurfactant from bacillus subtilis. *Biotechnology journal*, 10(8), 1216–1234.

- Danos, V., & Laneve, C. (2004). Formal molecular biology. *Theoretical Computer Science*, 325(1), 69–110.
- Davis, M. J., & Skodje, R. T. (1999). Geometric investigation of low-dimensional manifolds in systems approaching equilibrium. *The Journal of chemical physics*, 111(3), 859–874.
- Del Vecchio, D. (2007). Design and analysis of an activator-repressor clock in *e. coli*. In *2007 american control conference* (pp. 1589–1594).
- Desel, J. (1990). Reduction and design of well-behaved concurrent systems. In *International conference on concurrency theory* (pp. 166–181).
- Érdi, P., & Tóth, J. (1989). *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models*. Manchester University Press.
- Esparza, J. (1998). Decidability and complexity of petri net problems—an introduction. In *Lectures on petri nets i: Basic models* (pp. 374–428). Springer.
- Esparza, J., & Silva, M. (1990). Top-down synthesis of live and bounded free choice nets. In *International conference on application and theory of petri nets* (pp. 118–139).
- Fages, F., Gay, S., & Soliman, S. (2015). Inferring reaction systems from ordinary differential equations. *Theoretical Computer Science*, 599, 64–78.
- Fages, F., Soliman, S., & Chabrier-Rivier, N. (2004). Modelling and querying interaction networks in the biochemical abstract machine biocham. *Journal of Biological Physics and Chemistry*, 4, 64–73.
- Feinberg, M. (1987). Chemical reaction network structure and the stability of complex isothermal reactors—i. the deficiency zero and deficiency one theorems. *Chemical Engineering Science*, 42(10), 2229–2268.
- Feliu, E., & Wiuf, C. (2012). Variable elimination in chemical reaction networks with mass-action kinetics. *SIAM Journal on Applied Mathematics*, 72(4), 959–981.
- Fournet, C., & Gonthier, G. (1996). The reflexive cham and the join-calculus. In *Proceedings of the 23rd acm sigplan-sigact symposium on principles of programming languages* (pp. 372–385).
- Fraser, S. J. (1988). The steady state and equilibrium approximations: A geometrical picture. *The Journal of chemical physics*, 88(8), 4732–4738.
- Gagneur, J., & Klamt, S. (2004). Computation of elementary modes: a unifying framework and the new binary approach. *BMC bioinformatics*, 5(1), 175.
- Gay, S., Fages, F., Martinez, T., Soliman, S., & Solnon, C. (2014). On the subgraph epimorphism problem. *Discrete Applied Mathematics*, 162, 214–228.
- Gay, S., Soliman, S., & Fages, F. (2010). A graphical method for reducing and relating models in systems biology. *Bioinformatics*, 26(18), i575–i581.

- Genrich, H. J. (1988). Equivalence transformations of prt-nets. In *European workshop on applications and theory in petri nets* (pp. 179–208).
- Gibbs, J. W. (2014). *Elementary principles in statistical mechanics*. Courier Corporation.
- Gillespie, C. S. (2009). Moment-closure approximations for mass-action models. *IET systems biology*, 3(1), 52–58.
- Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25), 2340–2361.
- Gorban, A. N., & Karlin, I. V. (2003). Method of invariant manifold for chemical kinetics. *Chemical Engineering Science*, 58(21), 4751–4768.
- Gorban, A. N., Karlin, I. V., Ilg, P., & Öttinger, H. C. (2001). Corrections and enhancements of quasi-equilibrium states. *Journal of non-newtonian fluid mechanics*, 96(1), 203–219.
- Gossen, M., & Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences*, 89(12), 5547–5551.
- Gossen, M., Freundlieb, S., Bender, G., Müller, G., Hillen, W., & Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. *Science*, 268(5218), 1766.
- Gunawardena, J. (2012). A linear framework for time-scale separation in nonlinear biochemical systems. *PLoS One*, 7(5), e36321.
- Haddad, S. (1991). A reduction theory for coloured nets. In *High-level petri nets* (pp. 399–425). Springer.
- Haddad, S., & Pradat-Peyre, J.-F. (2006). New efficient petri nets reductions for parallel programs verification. *Parallel Processing Letters*, 16(01), 101–116.
- Heineken, F., Tsuchiya, H., & Aris, R. (1967). On the mathematical status of the pseudo-steady state hypothesis of biochemical kinetics. *Mathematical Biosciences*, 1(1), 95–113.
- Heiner, M., Gilbert, D., & Donaldson, R. (2008). Petri nets for systems and synthetic biology. In *International school on formal methods for the design of computer, communication and software systems* (pp. 215–264).
- Heiner, M., Koch, I., & Will, J. (2004). Model validation of biological pathways using petri nets—demonstrated for apoptosis. *Biosystems*, 75(1), 15–28.
- Heinrich, R., & Schuster, S. (2012). *The regulation of cellular systems*. Springer Science & Business Media.
- Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., ... Kummer, U. (2006). Copasi—a complex pathway simulator. *Bioinformatics*, 22(24), 3067–3074.
- Horiuti, J., & Nakamura, T. (1957). Stoichiometric number and the theory of steady reaction. *Z. Phys. Chem., Neue Folge*, 11, 358–365.

- Horn, F. (1972). Necessary and sufficient conditions for complex balancing in chemical kinetics. *Archive for Rational Mechanics and Analysis*, 49(3), 172–186.
- Horn, F., & Jackson, R. (1972). General mass action kinetics. *Archive for rational mechanics and analysis*, 47(2), 81–116.
- Huang, Z., Moya, C., Jayaraman, A., & Hahn, J. (2010). Using the tet-on system to develop a procedure for extracting transcription factor activation dynamics. *Molecular bioSystems*, 6(10), 1883–1889.
- Huet, G. (1980). Confluent reductions: Abstract properties and applications to term rewriting systems: Abstract properties and applications to term rewriting systems. *Journal of the ACM (JACM)*, 27(4), 797–821.
- John, M., Lhoussaine, C., Niehren, J., & Versari, C. (2011). Biochemical reaction rules with constraints. In *European symposium on programming* (pp. 338–357).
- Johnsonbaugh, R., & Murata, T. (1981). Additional methods for reduction and expansion of marked graphs. *IEEE Transactions on Circuits and Systems*, 28(10), 1009–1014.
- Juan, E. Y., Tsai, J. J., & Murata, T. (1998). Compositional verification of concurrent systems using petri-net-based condensation rules. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 20(5), 917–979.
- Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of theoretical biology*, 22(3), 437–467.
- Kauffman, S. A. (1993). *The origins of order: Self organization and selection in evolution*. Oxford University Press, USA.
- King, E. L., & Altman, C. (1956). A schematic method of deriving the rate laws for enzyme-catalyzed reactions. *The Journal of physical chemistry*, 60(10), 1375–1378.
- Kitano, H. (2002). Systems biology: a brief overview. *Science*, 295(5560), 1662–1664.
- Kosaraju, S. R. (1982). Decidability of reachability in vector addition systems (preliminary version). In *Proceedings of the fourteenth annual acm symposium on theory of computing* (pp. 267–281).
- Kovalyov, A. (1990). On complete reducibility of some classes of petri nets. In *Proceedings of the 11th international conference on applications and theory of petri nets* (pp. 352–366).
- Krohn, P., & Rauhamaa, M. (1988). *Reduction transformations of prt-nets* (Tech. Rep.). Helsinki University of Technology, Digital Systems Laboratory.
- Kron, G. (1939). *Tensor analysis of networks*. J. Wiley & Sons.
- Kuo-Chen, C., & Forsen, S. t. (1981). Graphical rules of steady-state reaction systems. *Canadian Journal of Chemistry*, 59(4), 737–755.
- Kuttler, C., Lhoussaine, C., & Niehren, J. (2007). A stochastic pi calculus for

- concurrent objects. In *International conference on algebraic biology* (pp. 232–246).
- Lam, S., & Goussis, D. (1994). The csp method for simplifying kinetics. *International Journal of Chemical Kinetics*, 26(4), 461–486.
- Lee, K.-H., & Favrel, J. (1985). Hierarchical reduction method for analysis and decomposition of petri nets. *IEEE Transactions on Systems, Man, and Cybernetics*(2), 272–280.
- Lee-Kwang, H., Favrel, J., & Baptiste, P. (1987). Generalized petri net reduction method. *IEEE transactions on systems, man, and cybernetics*, 17(2), 297–303.
- Leloup, J.-C., & Goldbeter, A. (2003). Toward a detailed computational model for the mammalian circadian clock. *Proceedings of the National Academy of Sciences*, 100(12), 7051–7056.
- Lipton, R. (1976). The reachability problem requires exponential space. *Research Report 62, Department of Computer Science, Yale University, New Haven, Connecticut*.
- Lu, T., Ju, Y., & Law, C. K. (2001). Complex csp for chemistry reduction and analysis. *Combustion and Flame*, 126(1), 1445–1455.
- Maas, U., & Pope, S. B. (1992). Simplifying chemical kinetics: intrinsic low-dimensional manifolds in composition space. *Combustion and flame*, 88(3), 239–264.
- Madelaine, G., Lhoussaine, C., & Niehren, J. (2014). Attractor equivalence: An observational semantics for reaction networks. In *International conference on formal methods in macro-biology* (pp. 82–101).
- Madelaine, G., Lhoussaine, C., & Niehren, J. (2015). Structural simplification of chemical reaction networks preserving deterministic semantics. In *International conference on computational methods in systems biology* (pp. 133–144).
- Madelaine, G., Lhoussaine, C., Niehren, J., & Tonello, E. (2016). Structural simplification of chemical reaction networks in partial steady states. *Biosystems*.
- Madelaine, G., Tonello, E., Lhoussaine, C., & Niehren, J. (2016a). Normalizing chemical reaction networks by confluent structural simplification. In *International conference on computational methods in systems biology* (pp. 201–215).
- Madelaine, G., Tonello, E., Lhoussaine, C., & Niehren, J. (2016b). *Simplification of reaction networks, confluence and elementary modes*. (Soumis)
- Mäder, U., Schmeisky, A. G., Flórez, L. A., & Stülke, J. (2011). Subtiwiki—a comprehensive community resource for the model organism bacillus subtilis. *Nucleic acids research*, gkr923.
- Markevich, N. I., Hoek, J. B., & Kholodenko, B. N. (2004). Signaling switches



- and bistability arising from multisite phosphorylation in protein kinase cascades. *The Journal of cell biology*, 164(3), 353–359.
- Masuda, M. (1990). Stoichiometric relations of complex chemical reaction network and stoichiometric number theory: A criterion for preserving overall stoichiometry in reduced mechanisms. *The Journal of Chemical Physics*, 92(10), 6030–6035.
- Matache, M. T., & Matache, V. (2016). Logical reduction of biological networks to their most determinative components. *Bulletin of Mathematical Biology*, 78(7), 1520–1545.
- Mayr, E. W. (1984). An algorithm for the general petri net reachability problem. *SIAM Journal on computing*, 13(3), 441–460.
- Michaelis, L., & Menten, M. L. (1913). Die kinetik der invertinwirkung. *Biochem. z*, 49(333-369), 352.
- Milner, R. (1989). *Communication and concurrency* (Vol. 84). Prentice hall New York etc.
- Mitchell, J. C. (1996). *Foundations for programming languages* (Vol. 1). MIT press Cambridge.
- Murata, T. (1989). Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, 77(4), 541–580.
- Murata, T., & Koh, J. (1980). Reduction and expansion of live and safe marked graphs. *IEEE transactions on circuits and systems*, 27(1), 68–71.
- Naldi, A., Monteiro, P. T., & Chaouiya, C. (2012). Efficient handling of large signalling-regulatory networks by focusing on their core control. In *Computational methods in systems biology* (pp. 288–306).
- Naldi, A., Remy, E., Thieffry, D., & Chaouiya, C. (2011). Dynamically consistent reduction of logical regulatory graphs. *Theoretical Computer Science*, 412(21), 2207–2218.
- Niehren, J. (2000). Uniform confluence in concurrent computation. *Journal of Functional Programming*, 10(05), 453–499.
- Norris, J. R. (1998). *Markov chains, cambridge series in statistical and probabilistic mathematics, vol. 2*. Cambridge University Press, Cambridge.
- Okino, M. S., & Mavrovouniotis, M. L. (1998). Simplification of mathematical models of chemical reaction systems. *Chemical reviews*, 98(2), 391–408.
- Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? *Nature biotechnology*, 28(3), 245–248.
- Palsson, B. O. (1987). On the dynamics of the irreversible michaelis-menten reaction mechanism. *Chemical engineering science*, 42(3), 447–458.
- Paulevé, L. (2016, September). Goal-oriented reduction of automata networks. In *14th International Conference on Computational Methods in Systems Biology (CMSB 2016)*.
- Petri, C. A. (1966). Communication with automata.

- Pitts, A. M. (2002). Operational semantics and program equivalence. In *Applied semantics* (pp. 378–412). Springer.
- Popper, K. (2005). *The logic of scientific discovery*. Routledge.
- Prigent, S., Collet, G., Dittami, S. M., Delage, L., Ethis de Corny, F., Dameron, O., ... Tonon, T. (2014). The genome-scale metabolic network of *ectocarpus siliculosus* (ectogem): a resource to study brown algal physiology and beyond. *The Plant Journal*, 80(2), 367–381.
- Purnick, P. E., & Weiss, R. (2009). The second wave of synthetic biology: from modules to systems. *Nature reviews Molecular cell biology*, 10(6), 410–422.
- Qu, Z., MacLellan, W. R., & Weiss, J. N. (2003). Dynamics of the cell cycle: checkpoints, sizers, and timers. *Biophysical journal*, 85(6), 3600–3611.
- Queille, J.-P., & Sifakis, J. (1982). Specification and verification of concurrent systems in cesar. In *International symposium on programming* (pp. 337–351).
- Radulescu, O., Gorban, A. N., Zinovyev, A., & Lilienbaum, A. (2008). Robust simplifications of multiscale biochemical networks. *BMC systems biology*, 2(1), 86.
- Radulescu, O., Gorban, A. N., Zinovyev, A., & Noel, V. (2012). Reduction of dynamical biochemical reaction networks in computational biology. *arXiv preprint arXiv:1205.2851*.
- Radulescu, O., Vakulenko, S., & Grigoriev, D. (2015). Model reduction of biochemical reactions networks by tropical analysis methods. *Mathematical Modelling of Natural Phenomena*, 10(3), 124–138.
- Rao, A., Hero III, A. O., & Engel, J. D. (2007). Using directed information to build biologically relevant influence networks. In *Comput syst bioinform conf (s1752-7791)* (Vol. 6, pp. 145–156).
- Rao, C. V., & Arkin, A. P. (2003). Stochastic chemical kinetics and the quasi-steady-state assumption: application to the gillespie algorithm. *The Journal of chemical physics*, 118(11), 4999–5010.
- Rao, S., Jayawardhana, B., & van der Schaft, A. (2016). Model reduction of detailed-balanced reaction networks by clustering linkage classes.
- Rao, S., Van der Schaft, A., Van Eunen, K., Bakker, B. M., & Jayawardhana, B. (2014). A model reduction method for biochemical reaction networks. *BMC systems biology*, 8(1), 1.
- Reddy, V. N., Mavrovouniotis, M. L., & Liebman, M. N. (1993). Petri net representations in metabolic pathways. In *Ismb* (Vol. 93, pp. 328–336).
- Regev, A., Panina, E. M., Silverman, W., Cardelli, L., & Shapiro, E. (2004). Bioambients: an abstraction for biological compartments. *Theoretical Computer Science*, 325(1), 141–167.

- Richardson, K. A. (2005). Simplifying boolean networks. *Advances in Complex Systems*, 8(04), 365–381.
- Roussel, M. R., & Fraser, S. J. (1990). Geometry of the steady-state approximation: Perturbation and accelerated convergence methods. *The Journal of chemical physics*, 93(2), 1072–1081.
- Saadatpour, A., Albert, I., & Albert, R. (2010). Attractor analysis of asynchronous boolean models of signal transduction networks. *Journal of theoretical biology*, 266(4), 641–656.
- Saadatpour, A., Albert, R., & Reluga, T. C. (2013). A reduction method for boolean network models proven to conserve attractors. *SIAM Journal on Applied Dynamical Systems*, 12(4), 1997–2011.
- Sáez, M., Wiuf, C., & Feliu, E. (2016). Graphical reduction of reaction networks by linear elimination of species. *Journal of Mathematical Biology*, 1–43.
- Sanft, K. R., Gillespie, D. T., & Petzold, L. R. (2011). Legitimacy of the stochastic michaelis & menten approximation. *IET systems biology*, 5(1), 58–69.
- Schmidt-Schauss, M., Sabel, D., Niehren, J., & Schwinghammer, J. (2015). Observational program calculi and the correctness of translations. *Theoretical Computer Science*, 577, 98–124.
- Schmierer, B., Tournier, A. L., Bates, P. A., & Hill, C. S. (2008). Mathematical modeling identifies smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system. *Proceedings of the National Academy of Sciences*, 105(18), 6608–6613.
- Schnell, S., & Mendoza, C. (1997). Closed form solution for time-dependent enzyme kinetics. *Journal of theoretical Biology*, 187(2), 207–212.
- Schnoebelen, P., & Sidorova, N. (2000). Bisimulation and the reduction of petri nets. In *International conference on application and theory of petri nets* (pp. 409–423).
- Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D., & Müller, G. (2002). Computational modeling of the dynamics of the map kinase cascade activated by surface and internalized egf receptors. *Nature biotechnology*, 20(4), 370–375.
- Schuster, S., & Hilgetag, C. (1994). On elementary flux modes in biochemical reaction systems at steady state. *Journal of Biological Systems*, 2(02), 165–182.
- Segel, L. A. (1988). On the validity of the steady state assumption of enzyme kinetics. *Bulletin of mathematical biology*, 50(6), 579–593.
- Segel, L. A., & Slemrod, M. (1989). The quasi-steady-state assumption: a case study in perturbation. *SIAM review*, 31(3), 446–477.
- Shatz, S. M., Tu, S., Murata, T., & Duri, S. (1996). An application of petri net reduction for ada tasking deadlock analysis. *IEEE Transactions on Parallel and Distributed Systems*, 7(12), 1307–1322.

- Silva, M. (2013). Half a century after carl adam petri's ph. d. thesis: A perspective on the field. *Annual reviews in control*, 37(2), 191–219.
- Silva, M., & Murata, T. (1992). B-fairness and structural b-fairness in petri net models of concurrent systems. *Journal of Computer and System Sciences*, 44(3), 447–477.
- Sloan, R. H., & Buy, U. (1996). Reduction rules for time petri nets. *Acta Informatica*, 33(7), 687–706.
- Stiefenhofer, M. (1998). Quasi-steady-state approximation for chemical reaction networks. *Journal of Mathematical Biology*, 36(6), 593–609.
- Thiele, I., Swainston, N., Fleming, R. M., Hoppe, A., Sahoo, S., Aurich, M. K., . . . Palsson, B. (2013). A community-driven global reconstruction of human metabolism. *Nature biotechnology*, 31(5), 419–425.
- Thomas, R. (1973). Boolean formalization of genetic control circuits. *Journal of theoretical biology*, 42(3), 563–585.
- Thomas, R. (1991). Regulatory networks seen as asynchronous automata: a logical description. *Journal of theoretical biology*, 153(1), 1–23.
- Turanyi, T., Tomlin, A., & Pilling, M. (1993). On the error of the quasi-steady-state approximation. *The Journal of Physical Chemistry*, 97(1), 163–172.
- Uhendorf, J., Miermont, A., Delaveau, T., Charvin, G., Fages, F., Bottani, S., . . . Hersen, P. (2012). Long-term model predictive control of gene expression at the population and single-cell levels. *Proceedings of the National Academy of Sciences*, 109(35), 14271–14276.
- Veliz-Cuba, A. (2011). Reduction of boolean network models. *Journal of theoretical biology*, 289, 167–172.
- Veliz-Cuba, A., Laubenbacher, R., & Aguilar, B. (2013). Dimension reduction of large and-not network models. *arXiv preprint arXiv:1311.6868*.
- Volkenstein, M. V., & Goldstein, B. N. (1966). A new method for solving the problems of the stationary kinetics of enzymological reactions. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 115(2), 471–477.
- Winkel, G. (1993). *The formal semantics of programming languages: an introduction*. MIT press.
- Wolf, K., Stahl, C., Ott, J., & Danitz, R. (2009). Verifying deadlock-and livelock freedom in an soa scenario. In *Application of concurrency to system design, 2009. acsd'09. ninth international conference on* (pp. 168–177).
- Wong, J. T.-F., & Hanes, C. S. (1962). Kinetic formulations for enzymic reactions involving two substrates. *Canadian journal of biochemistry and physiology*, 40(6), 763–804.
- Zagaris, A., Kaper, H. G., & Kaper, T. J. (2004). Analysis of the computational singular perturbation reduction method for chemical kinetics. *Journal of Nonlinear Science*, 14(1), 59–91.
- Zañudo, J. G., & Albert, R. (2013). An effective network reduction approach

- to find the dynamical repertoire of discrete dynamic networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(2), 025111.
- Zevedei-Oancea, I., & Schuster, S. (2003). Topological analysis of metabolic networks based on petri net theory. *In silico biology*, 3(3), 323–345.
- Zhou, G., & Deng, M. (1984). An extension of chou's graphic rules for deriving enzyme kinetic equations to systems involving parallel reaction pathways. *Biochemical Journal*, 222(1), 169–176.

## Résumé

La biologie des systèmes cherche à comprendre et analyser des systèmes biologiques à l'aide de modèles mathématiques et informatiques. L'explosion des données expérimentales entraîne des modèles de plus en plus grands. Afin de pouvoir les analyser facilement ou les simuler rapidement, il est alors nécessaire de pouvoir les simplifier.

Dans cette thèse, nous proposons des méthodes de simplifications de réseaux de réactions biochimiques. Ces méthodes sont suffisamment riches pour pouvoir simplifier un grand nombre de réseaux provenant d'applications biologiques. Elles sont contextuelles, permettant de considérer un réseau comme un sous-module d'un modèle plus grand, et de le simplifier sans modifier le comportement du modèle global. Enfin, nos simplifications sont correctes, c'est-à-dire qu'elles préservent la sémantique des réseaux. Dans un premier temps, nous nous intéressons à une sémantique non déterministe, basée sur la possibilité de converger vers les attracteurs du réseau. Nous étudions ensuite des simplifications pour la sémantique déterministe, permettant par exemple de supprimer des espèces intermédiaires à l'équilibre. Enfin, nous nous intéressons à la confluence d'une telle simplification, ainsi qu'au lien entre l'élimination des espèces intermédiaires et le calcul des modes élémentaires d'un réseau.

**Mots clés :** biologie des systèmes; réseaux de réactions chimique; réseaux de Pétri; simplification; règle de réécriture; équivalence; contexte; sémantique non-déterministe; sémantique observationnelle; attracteurs; convergence; sémantique déterministe; équations différentielles ordinaires; état stable; espèces intermédiaires; confluence; modes élémentaires.

## Abstract

System biology aims at understanding and analyzing biological systems using mathematical and computational models. The explosion of the number of experimental data leads to larger and larger models. In order to be able to easily analyze them and quickly simulate them, it is necessary to be able to simplify them.

In this thesis, we propose simplification methods for biochemical reaction networks. These methods are sufficiently rich to be able to simplify an important number of networks from biological applications. They are contextual, allowing to consider a network as a sub-module of a larger model, and to simplify it without modifying the behavior of the global model. Finally, our simplifications are sound, meaning that they preserve the semantics of the networks. Firstly, we are interested in a non deterministic semantics, based on the capability to converge to some attractors of the network. Then we study some simplifications for the deterministic semantics, allowing for instance to remove intermediate species at steady-state. Finally, we are interested by the confluence of this simplification, as well as the relation between the elimination of intermediate species and the computation of the elementary modes of a network.

**Keywords:** system biology; chemical reaction networks; Petri nets; simplification; rewriting rules; equivalence; context; non deterministic semantics; observational semantics; attractors; convergence; deterministic semantics; ordinary differential equations; steady state; intermediate species; confluence; elementary modes.