

**Université de Montréal**

**Sur l'application de la structure de graphes pour le  
calcul automatique de nombres de reproduction dans  
les modèles à compartiments déterministes**

par

**Alexandre Simard**

Département d'informatique et de recherche opérationnelle  
Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de  
Maître ès sciences (M. Sc.)  
en informatique

Orientation intelligence artificielle

30 avril 2023



# Université de Montréal

Faculté des arts et des sciences

Ce mémoire intitulé

**Sur l'application de la structure de graphes pour  
le calcul automatique de nombres de reproduction  
dans les modèles à compartiments déterministes**

présenté par

**Alexandre Simard**

a été évalué par un jury composé des personnes suivantes :

*Gena Hahn*

---

(président-rapporteur)

*Irina Rish*

---

(directeur de recherche)

*Jacques Bélair*

---

(codirecteur)

*Jean-Yves Potvin*

---

(membre du jury)



# Résumé

---

En basant l'analyse des modèles épidémiologiques sur leur représentation graphique plutôt que sur leurs équations différentielles, il est possible de mettre en évidence plusieurs concepts importants à l'aide des composantes d'un hypergraphe. On décrit une manière formelle de créer automatiquement un système d'équations différentielles à partir de ces composantes et on adapte ensuite la définition du produit cartésien [25] pour les hypergraphes décrits, ce qui permet la fusion de modèles.

À l'aide d'un algorithme qui ajoute automatiquement de nouvelles composantes à l'hypergraphe, il est possible d'isoler virtuellement certains individus, afin d'explicitier le calcul de nombres de reproduction [17]. On montre ensuite que la forme des équations différentielles créées admettent une solution unique [14] et que l'algorithme d'ajout aux hypergraphes est stable au niveau de la structure et de la dynamique des hypergraphes.

On trouve que la méthode décrite pour le calcul des nombres de reproduction permet une meilleure prédiction de la croissance de l'épidémie que le calcul standard  $\mathcal{R}_t = \mathcal{R}_0 \cdot S/N$  [12] et que le calcul de  $\mathcal{R}_0$  est très similaire aux résultats trouvés à l'aide de la matrice de prochaine génération [6, 23], en plus d'être plus simple à mettre en place et d'offrir une justification plus robuste.

On conclue ce mémoire en décrivant sommairement un processus d'apprentissage automatique des paramètres dans les modèles à compartiments, afin de permettre une calibration de modèles plus rapide. L'apprentissage machine peut être intégré en faisant appel à la librairie `torchdiffeq` [4], qui implémente les équations différentielles ordinaires neuronales [5] en utilisant Pytorch.

**Mots clés** : Épidémiologie mathématique; Modèles déterministes; Théorie des graphes; Équations différentielles; Nombres de reproduction; Apprentissage automatique.



# Abstract

---

By basing the analysis of epidemiological models on their graphical representation rather than on their differential equations, it is possible to highlight a few key concepts by using the components of a hypergraph. We give a formal way to automatically create a system of differential equations by using these components and we then adapt the definition of the cartesian product [25] for the defined hypergraphs, which permits the merging of models.

Using an algorithm which automatically adds new components to the graph, we can virtually isolate a few individuals to explicitly compute the reproduction numbers [17]. We then show that the resulting differential equations allow for a unique solution [14] and that the modification algorithm is stable for the structure and dynamics of the hypergraphs.

We find that the described method for the computation of reproduction numbers gives a more accurate prediction of the growth of the epidemic than the standard computation  $\mathcal{R}_t = \mathcal{R}_0 \cdot S/N$  [12] and that the computation of  $\mathcal{R}_0$  is very similar to the results found using the next generation matrix method [6, 23], as well as being simpler to integrate into models and offering a more robust justification.

We conclude this thesis with a brief outline of an automatic learning process for the parameters in compartmental models, which allows a faster calibration of epidemiological models. The implementation of machine learning can be done through the torchdiffeq library [4], which applies the theory of neural ordinary differential equations [5] using Pytorch.

**Mots clés:** Mathematical epidemiology; Deterministic models; Graph theory; Differential equations; Reproduction numbers; Machine learning.



# Table des matières

---

<b>Résumé</b> .....	5
<b>Abstract</b> .....	7
<b>Liste des tableaux</b> .....	11
<b>Table des figures</b> .....	13
<b>Liste des sigles et des abréviations</b> .....	15
<b>Remerciements</b> .....	19
<b>Chapitre 1. Introduction</b> .....	21
<b>Chapitre 2. Formalisation des modèles épidémiologiques</b> .....	25
2.1. Carquois épidémiologiques .....	28
2.2. Carquois épidémiologiques .....	30
2.3. Système d'équations différentielles .....	32
2.4. Produits d'hypergraphes .....	34
2.5. Structure informatique .....	36
<b>Chapitre 3. Calcul explicite de <math>\mathcal{R}_0</math> et <math>\mathcal{R}_t</math></b> .....	39
3.1. Algorithme d'ajout aux hypergraphes .....	41
3.2. Conditions initiales du modèle modifié .....	45
3.3. Sur la normalisation de $\mathcal{R}_t$ .....	47
3.4. Complexité des algorithmes .....	48

<b>Chapitre 4. Propositions et preuves</b> .....	51
4.1. Unicité des solutions .....	51
4.2. Structure du graphe modifié .....	53
4.3. Dynamique du modèle modifié .....	55
4.4. Normalisation du nombre de cas index .....	58
4.5. Convergence des nœuds $\mathcal{R}_t$ .....	59
<b>Chapitre 5. Capacité maximale</b> .....	61
5.1. Deux types de capacité .....	62
5.2. Erratum : section 2.3 .....	64
5.3. Capacité sur population .....	65
5.4. Capacité du modèle modifié .....	66
<b>Chapitre 6. Résultats</b> .....	67
6.1. Modèle <i>SIR</i> .....	67
6.2. Modèles plus complexes .....	69
6.3. Maximum d'incidence totale .....	73
<b>Conclusion</b> .....	81
Apprentissage automatique des paramètres .....	82
<b>Références bibliographiques</b> .....	87
<b>Annexe A. Modèles additionnels et modifications</b> .....	89
<b>Annexe B. Quelques pseudocodes</b> .....	93

## Liste des tableaux

---

1.1	Définition de plusieurs variables utilisés en épidémiologie. ....	23
3.1	Nouveaux hyperarcs pour le modèle de la figure 3.4. ....	45
5.1	Exemples de fonctions pour représenter une capacité stricte. ....	63
5.2	Exemples de fonctions pour représenter une capacité souple. ....	63
6.1	Convergence de $\mathcal{R}_0^{\text{sim}}$ selon la séparation du compartiment $I$ . Modèle $SIR$ simple, $\beta = 0.8, \gamma = 0.2$ . ....	70



## Table des figures

---

2.1	Modèle <i>SIRS</i> à compartiments, où les arêtes sont mal définies. ....	26
2.2	arêtes bien définies. ....	27
2.3	Exemple d'un graphe orienté. ....	28
2.4	Exemple d'un hypergraphe. ....	29
2.5	Simplification de l'hypergraphe donné à l'exemple 2.1.8. ....	31
2.6	Figure 2.5 avec l'inclusion des nœuds Null. ....	32
2.7	Produit de deux hypergraphes $G$ et $H$ . ....	35
3.1	Modèle <i>SEIR</i> simple. ....	41
3.2	Modèle <i>SIRS</i> avec démographie et vaccination. ....	43
3.3	Modification du modèle <i>SIRS</i> de la figure 3.2. ....	43
3.4	Modèle <i>SIS</i> avec démographie. ....	44
3.5	Modification d'un modèle <i>SIS</i> avec démographie (figure 3.4). ....	45
5.1	Modèle <i>SIRS</i> avec hospitalisation des infectés. ....	61
5.2	Courbes d'infection pour un modèle <i>SIRS</i> avec hospitalisation. ....	62
6.1	Modèle <i>SIR</i> simple. ....	67
6.2	Calcul de $\mathcal{R}_t$ pour deux valeurs de paramètres. ....	68
6.3	Courbes $\mathcal{R}_t$ pour le modèle de la figure A.3. ....	69
6.4	Courbes $\mathcal{R}_t$ pour deux modèles avec phase endémique. ....	70
6.5	Courbes $\mathcal{R}_t$ pour le modèle de la figure A.9 et le <i>SIR</i> de base associé. ....	71

6.6	Effet de changements soudains des paramètres sur les courbes de reproduction. La valeur de $\mathcal{R}_0$ change de 4 à 3 au temps $t = 10$ , puis de 3 à 5 au temps $t = 40$ . . . .	72
6.7	Courbes $\mathcal{R}_t$ pour le modèle de la figure A.12. . . . .	73
6.8	Valeur de $ \mathcal{R}_{t^*}^{\text{sim}} - 1 $ pour un ensemble de valeurs de $\beta$ et $\gamma$ entre 0 et 10 telles que $\mathcal{R}_0 > 1$ , lorsque $N = 10^6$ . . . . .	79
6.9	Modèle de prédiction pour les nombres de reproduction à partir d'une condition initiale et des paramètres $f$ . . . . .	83
6.10	Modèle de prédiction pour les nombres de reproduction à partir d'une condition initiale et des paramètres $f$ , pour la méthode proposée. . . . .	84
A.1	Modèle <i>SIS</i> avec vaccination. . . . .	89
A.2	Modification du modèle de la figure A.1. . . . .	89
A.3	Modèle <i>SIR</i> avec deux étapes d'infection. . . . .	89
A.4	Modification du modèle de la figure A.3. . . . .	90
A.5	Modèle <i>SIRS</i> avec démographie et vaccination. . . . .	90
A.6	Modification du modèle de la figure A.5. . . . .	90
A.7	Modèle <i>SIR</i> avec mortalité et vaccination. . . . .	90
A.8	Modification du modèle de la figure A.7. . . . .	91
A.9	Modèle <i>SIR</i> avec séparation séquentielle des infectés. . . . .	91
A.10	Modification du modèle de la figure A.9. . . . .	91
A.11	Modification du modèle <i>SIR</i> simple (figure 6.1). . . . .	91
A.12	Modèle <i>SIRS</i> avec taux de contacts périodiques. . . . .	91

## Liste des sigles et des abréviations

---

COVID-19	La maladie à coronavirus 2019
EDOs	Équations différentielles ordinaires
INSPQ	Institut national de santé publique du Québec
SEIR	Modèle Susceptibles → Exposés → Infectés → Rétablis
SEIS	Modèle Susceptibles → Exposés → Infectés → Susceptibles
SIR	Modèle Susceptibles → Infectés → Rétablis
SIRS	Modèle Susceptibles → Infectés → Rétablis → Susceptibles
SIS	Modèle Susceptibles → Infectés → Susceptibles



*Pour vrai live, je tiens pas à grand chose.*

Marc-Antoine Trahan,  
*Compétition MCM, 2021*

*... and the places of the dwelling of Ilúvatar were filled to overflowing,  
and the music and the echo of the music went out into the Void,  
and it was not void.*

J. R. R. Tolkien,  
*The Silmarillion*



# Remerciements

---

Je tiens d'abord à remercier ma directrice de recherche Irina Rish et mon codirecteur Jacques Bélair pour leur encadrement dans ce projet. Je remercie M. Bélair pour avoir accepté de commencer la supervision de ce projet alors que ma maîtrise était depuis longtemps entamée. Je vous suis reconnaissant pour vos conseils qui ont permis d'amener ce projet à bons termes. Je remercie également M<sup>me</sup> Rish pour son soutien financier et son appui durant le changement de directeur.

Je veux également remercier mon précédent directeur de recherche Simon de Montigny pour m'avoir aidé avec le commencement de ce projet et sa mise en forme, ainsi que pour son support financier. C'est vous qui avez eu l'idée initiale pour ce projet et qui m'avez donné l'opportunité de travailler sur une application aussi importante.

Je n'oublie pas les professeur·e·s du département de mathématiques et de statistique, qui m'ont transmis leurs connaissances tout au long de mon parcours universitaire et qui m'ont introduit au monde de la modélisation. En particulier, Yvan Saint-Aubin m'a introduit à la compétition de modélisation MCM, sans laquelle je n'aurais pas accompli cette maîtrise, et Morgan Craig m'a offert un support incomparable.

Je remercie Shophika de m'avoir convaincu d'écrire à Simon pour faire cette maîtrise. Je remercie Blanche, Jeanne, Francis, Laurent et Irene qui m'ont aidé à la rédaction de ce mémoire en m'écoutant parler de mes diverses impasses et en relisant certaines sections. Vos critiques et conseils ont grandement amélioré ce projet.

Un grand merci à ma famille, à mes ami·e·s et à ma chère Marie, qui m'ont soutenu au cours de mes études.



# Chapitre 1

---

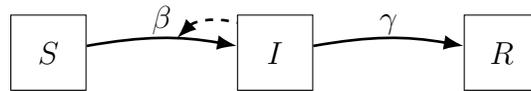
## Introduction

Lorsqu'on se trouve en contexte d'épidémie, l'épidémiologie mathématique est cruciale pour étudier la dynamique de la maladie à travers un ensemble d'outils qui sont formels et généralisables, et qui offrent un ensemble de conclusions réfutables [17]. Par exemple, les modèles à compartiments tels que le *SIR* ou *SIRS*, décrits par un système d'EDOs, sont utilisés pour calculer les nombres de reproduction  $\mathcal{R}_0$  et  $\mathcal{R}_t$  associés à une maladie [8]. Le nombre de reproduction de base  $\mathcal{R}_0$  est défini comme le nombre d'infections secondaires que produira un individu infecté durant sa période d'infection, s'il est placé dans une population complètement susceptible [23]. Le nombre de reproduction effectif  $\mathcal{R}_t$  représente quant à lui le nombre d'infections produites si l'individu infecté à l'étude est introduit au temps  $t$ , là où une certaine immunité s'est développée dans la population [9].

Bien que ces quantités soient hautement importantes pour l'étude de maladies infectieuses, les méthodes existantes pour le calcul des nombres de reproduction dans les modèles à compartiments sont difficiles à mettre en place. Par exemple, la méthode de la matrice de prochaine génération pour le calcul de  $\mathcal{R}_0$  découle d'une analyse mathématique qui dépend grandement de la structure du modèle [6, 23] et sa mise en œuvre devient plus complexe lorsque le nombre de compartiments infectés devient plus grand. Par conséquent, l'utilisation de cette méthode nécessite des personnes qui connaissent bien la théorie des systèmes dynamiques. Pour comparer différentes structures de modèles, ceci n'est pas idéal puisqu'il faut reprendre l'analyse mathématique pour chaque modèle de manière indépendante.

On souhaite dans ce mémoire décrire une nouvelle méthode de calcul des nombres de reproduction qui ne repose pas directement sur la structure du modèle. Habituellement, la

visualisation d'un modèle épidémiologique se fait à travers un graphe où les nœuds sont les compartiments du modèle et les flèches sont les mouvements de la population entre ces compartiments. En basant l'analyse d'un modèle épidémiologique sur cette représentation graphique, par exemple :



qui représente un modèle dit « *SIR* » illustrant le parcours des individus susceptibles (*S*) qui deviennent infectés (*I*) puis remis (*R*) de cette infection, plutôt que sur ses équations différentielles,

$$\begin{aligned}
 S' &= -\beta S \frac{I}{N} \\
 I' &= \beta S \frac{I}{N} - \gamma I \\
 R' &= \gamma I,
 \end{aligned}$$

il est alors possible de mettre en évidence plusieurs concepts importants en utilisant les composantes de cette structure. À l'aide d'un algorithme qui ajoutera automatiquement de nouvelles composantes à celles existantes, les nombres de reproduction seront calculés explicitement dans une simulation du modèle basée sur une nouvelle représentation graphique.

Ceci permettra idéalement aux modélisateurs d'obtenir des informations pertinentes sur les modèles développés avec plus de facilité. Dans une optique d'explicabilité, il sera également plus transparent d'utiliser une méthode de calcul pour les nombres de reproduction qui est explicite pour alléger les communications entre différentes équipes de modélisation [3].

Suite à la formulation de cette méthode de calcul, le but premier sera de montrer que la structure de modèles épidémiologiques obtenue avec les graphes est bien définie. On montre entre autres que la dynamique du graphe se comporte comme attendu et que l'algorithme d'ajout de composantes ne modifie pas cette dynamique [14]. Puis, on étudie au chapitre 5 une manière d'implémenter la capacité maximale [22] à l'intérieur de la structure de modèles définie. Le chapitre 6 de ce mémoire sert finalement à comparer la méthode de calcul traditionnelle [9] à celle présentée. On y présente quelques modèles plus complexes ainsi qu'une analyse mathématique d'un résultat important sur le pic d'infection.

On formule sommairement à la fin de la conclusion un algorithme pour apprendre automatiquement les paramètres du modèle, qui se base sur le calcul automatique des nombres de reproduction et la library `torchdiffeq` [4].

On présente au tableau 1.1 le nom de quelques symboles couramment utilisés en épidémiologie, ainsi que leur définition au sens épidémiologique [17]. Une implémentation en Python de plusieurs notions présentées dans ce mémoire sera disponible sur GitHub<sup>1</sup> et une revue de la littérature sera présentée plus en détail au début des chapitres 2 et 3.

Symbole	Signification	Définition
$S$	Individus susceptibles	Population qui n'est pas présentement infectée et qui est à risque de contracter la maladie suite à des contacts.
$E$	Individus exposés	Population qui est présentement infectée par la maladie mais qui n'est pas encore contagieuse.
$I$	Individus infectés	Population qui est présentement infectée et qui est en mesure de transmettre la maladie.
$R$	Individus rétablis	Population qui présente une immunité, soit par exposition antérieure ou par vaccination.
$N$	Population totale	Population à l'étude pour le modèle en question.
$\beta$	Taux de contact	Le nombre de contacts de chaque personne par unité de temps multiplié par la probabilité que chaque contact mène à la transmission de la maladie.
$\gamma$	Taux de rétablissement	La probabilité qu'un individu infecté sorte du compartiment $I$ à chaque unité de temps.
$1/\gamma$	Durée d'infection	Temps moyen d'une période d'infection à partir du moment de contraction.
$\mu$	Taux de mortalité	La probabilité que chaque individu décède à chaque unité de temps.
$\Lambda$	Nombre de naissances	Le nombre de naissances à chaque unité de temps.
$\mathcal{R}_0$	Nombre de reproduction de base	Le nombre d'infections secondaires produites en moyenne lorsqu'une infection est introduite dans une population complètement susceptible.
$\mathcal{R}_t$	Nombre de reproduction effectif	Le nombre d'infections secondaires produites en moyenne lorsqu'une infection est introduite dans la population au temps $t$ .

TABLEAU 1.1 – Définition de plusieurs variables utilisés en épidémiologie.

1. <https://github.com/asimard1/AutomaticR0>



# Chapitre 2

---

## Formalisation des modèles épidémiologiques

En 2017, Worden et Porco [25] présentent une définition des modèles épidémiques en utilisant une structure de graphes et expliquent l'utilité de différents produits sur ceux-ci afin de combiner différentes modélisations de dynamiques dans la population. Leur approche présente les termes des équations différentielles sur les arêtes, sans toutefois montrer la construction du graphe à partir des équations, ou vice-versa. Le graphe est donc simplement un outil pour permettre la fusion de modèles, qu'il faut construire manuellement.

Worden et Porco présentent également la structure de manière mathématique, mais aucune implémentation informatique n'est montrée. Puisque les produits de graphes peuvent facilement atteindre des tailles qui entravent le travail à la main, il est difficile de voir comment implémenter à la fois la création de ces graphes, le calcul de produits entre deux modèles et l'écriture des nouvelles équations différentielles de manière efficace, pour ne pas limiter l'étude de plusieurs modélisation différentes.

L'étude des épidémies à travers des systèmes d'équations différentielles fait apparaître trois catégories de mouvements dans les équations :

- 1) Les termes sans proportionnalité : les mouvements qui ne dépendent pas de la taille des compartiments ni du nombre d'individus. Ces termes sont sous la forme  $f(t)$  ;
- 2) Les termes de taux : les mouvements qui dépendent de la taille d'une partie de la population, sans cependant être liés à des infections. Ils prennent la forme  $f(t)x$  ;
- 3) les termes d'infection : les mouvements qui dépendent de la taille de deux parties de la population, où une partie influence le mouvement de l'autre. Ces termes prennent la forme  $f(t)xy/N$ .

Dans ces termes,  $x$  joue un rôle de « proportionnalité » et  $y$  joue un rôle « d'infection ». Par exemple, dans le modèle  $SIR$  de base, on retrouve les termes  $\beta SI/N$  et  $\gamma I$  qui donnent respectivement la création des infections et le rétablissement des infectés. Dans le premier,  $S$  est de proportionnalité et  $I$  est d'infection, alors que dans le deuxième  $I$  est de proportionnalité. Les tailles de population qui apparaissent dans les termes d'équations différentielles peuvent aussi être des sommes. Par exemple, pour un terme créant des naissances à partir de la population complète, on aura le terme  $\mu(S + I + R)$ . Pour un terme d'infection où deux groupes influencent la vitesse de propagation, on aura plutôt un terme sous la forme  $\beta S(I_1 + I_2)/N$ , où  $S$  est de proportionnalité et  $(I_1 + I_2)$  est d'infection.

Le but de ce chapitre sera de construire une structure d'hypergraphes semblable à la structure de graphe présentée par Worden et Porco, qui permettra de créer le système d'équations différentielles directement à partir des arêtes. De cette manière, on crée un lien plus étroit entre la visualisation du modèle et ses équations mathématiques sous-jacentes, et on peut résoudre les limitations présentes dans la structure de graphe habituelle pour les modèles.

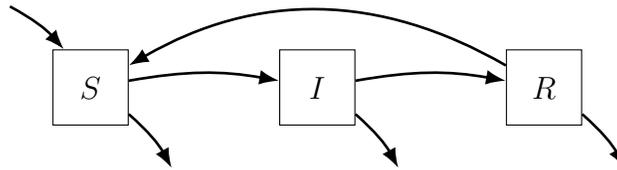


FIGURE 2.1 – Modèle  $SIRS$  à compartiments, où les arêtes sont mal définies.

Par exemple, si on souhaite étudier un modèle avec la structure de mouvements illustrée dans la figure 2.1, on voit que plusieurs systèmes d'équations différentielles différents peuvent correspondre à cette structure, même en faisant abstraction des valeurs des paramètres. En effet, il n'est pas clair avec ce schéma si :

- 1) les entrées sont proportionnelles à toute la population, proportionnelles seulement à  $S$ , ou simplement un nombre constant par jour (voir catégories de termes) ;
- 2) le transfert de  $S$  à  $I$  est réellement une infection, ou c'est plutôt celui entre  $I$  et  $R$  qui en est une (ici, le nom des compartiments donne un indice, mais cet indice pourrait disparaître avec une appellation différente) ;
- 3) l'infection est causée par le compartiment  $I$  ou par le compartiment  $R$ . (Le nom des compartiments donne encore un indice qui n'est pas universel. Par exemple, on

peut considérer un modèle *SEIS* avec une structure semblable, où l'infection serait toujours un mouvement entre les deux premiers compartiments, causé cependant par le compartiment à la droite.)

Pour représenter explicitement les infections dans la représentation de modèles à compartiments, on ajoute habituellement une flèche segmentée entre le compartiment qui cause l'infection et la flèche qui représente le mouvement des infections. Cependant, cela ne permet pas la représentation des termes de proportionnalité.

Pour régler ce problème, on montrera le terme de proportionnalité associé au terme d'une arête par une flèche pointillée reliant ce terme à l'arête et le terme d'infection par une flèche segmentée (on omettra ces flèches si ces termes n'existent pas). Par exemple, les arêtes montrées à la figure 2.2 ne peuvent définir les équations différentielles que d'une seule manière (en faisant abstraction des constantes). En particulier, on obtient pour  $S'$  une équation sous la forme  $S' = \alpha S - \beta SI/N$  et pour  $I'$  on obtient  $I' = \beta SI/N - \gamma I$ , où  $\alpha$ ,  $\beta$  et  $\gamma$  sont des paramètres réels quelconques.

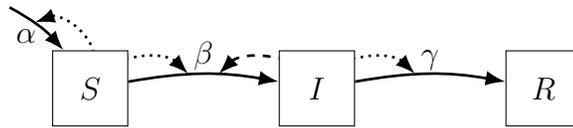


FIGURE 2.2 – arêtes bien définies.

Tel que mentionné, il sera question dans la création des équations différentielles de prendre des sommes, des produits et des quotients des compartiments du graphe. Si on ne donne aucune restriction sur les nœuds du graphe, il n'est pas assuré que ces opérations auront un sens. Dans le cas des modèles épidémiologiques présentés, les nœuds du graphe seront toujours des fonctions du temps et on n'aura pas ce problème.

Pour contourner ce problème dans la définition des structures utilisées (définition 2.1.7), on donne comme condition sur le graphe que les nœuds doivent faire partie d'un espace vectoriel. De cette manière, il sera possible d'évoquer la somme des nœuds d'une manière bien définie. Il est à noter que ces sommes (les termes de proportionnalité et d'infection) ne feront pas nécessairement partie du graphe lui-même. Par exemple, un terme de naissance  $\mu(S + I + R)$  dans un modèle *SIR* ferait appel à un élément extérieur au graphe.

## 2.1. Carquois épidémiologiques

On rappelle et définit dans cette section quelques notations et définitions se rapportant à la théorie des graphes, qu'on adapte à ce mémoire. On généralise ensuite ces définitions pour une représentation bien définie d'épidémies.

**Définition 2.1.1.** Un graphe orienté (ou digraphe)  $G$  est un couple  $G = (V, A)$  où  $V$  est un ensemble de nœuds et  $A$  est un ensemble d'arêtes. On aura  $A \subseteq \{(u, v) \in V^2 : u \neq v\}$  qui est un ensemble de couples orientés.

*Exemple 2.1.2.* On montre à la figure 2.3 un graphe orienté où

$$V(G) = \{1, 2, 3, 4\} \quad \text{et} \quad A(G) = \{(1, 2), (1, 4), (2, 3), (2, 4), (3, 4), (4, 1)\}.$$

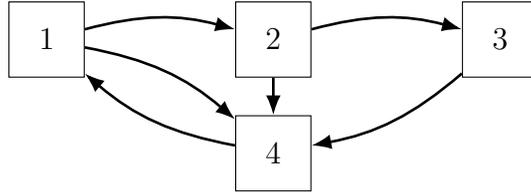


FIGURE 2.3 – Exemple d'un graphe orienté.

**Définition 2.1.3.** Un graphe de mouvement est un 3-tuple  $G = (V, A, f)$  où  $(V, A)$  est un graphe orienté et  $f : A \rightarrow \mathbb{R}^+$  est un paramètre de mouvement.

**Remarque 2.1.4.** Dans les schémas des graphes de mouvement, les paramètres de mouvement ne sont pas toujours représentés. On explicitera parfois ces paramètres en inscrivant leur valeur à proximité des arêtes.

**Définition 2.1.5.** Soit  $V$  un espace vectoriel et  $W$  un sous-ensemble de  $V$ . L'ensemble de toutes les sommes d'éléments de  $W$  est donné en utilisant l'ensemble des parties  $U$  de  $W$  :

$$\Sigma(W) = \left\{ \sum_{u \in U} u \in V : U \subseteq W \right\}. \quad (2.1.1)$$

*Notation 2.1.6.* Soit  $\sigma = \sum_{u \in U} u$ . On dira «  $v$  est dans la somme  $\sigma$  » si  $v \in U$ .

**Définition 2.1.7.** Un hypergraphe (aussi appelé carquois) est un couple  $G = (V, A)$  où  $V$  est un sous-ensemble fini d'un espace vectoriel et où chaque hyperarc dans  $A$  est un quadruplet. L'ensemble  $A$  est tel que

$$A \subseteq \{(u, v, v_r, v_c) \in V^2 \times (\Sigma(V))^2 : u \neq v\} = \mathcal{A}(V). \quad (2.1.2)$$

*Exemple 2.1.8.* Soit  $G = (V, A, f)$  avec  $V = \{a, b, c, d\}$  et

$$A = \{(a, b, a, b), (b, a, b, d), (c, a, a + b, d), (c, b, b, d), (a, d, a, d), (b, d, b, d)\}.$$

On montre un schéma pour ce carquois à la figure 2.4.

Comme on peut voir, ce genre de représentation devient rapidement difficile à suivre. On aura une manière de simplifier la représentation avec l'introduction de nœuds vides. Voir les définitions 2.2.8 et 2.2.3 ainsi que la remarque 2.2.12.

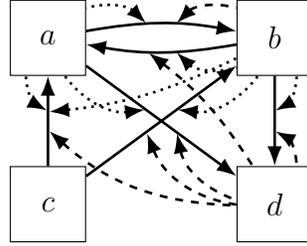


FIGURE 2.4 – Exemple d'un hypergraphe.

**Remarque 2.1.9.** On verra qu'inclure l'hyperarc  $(u, v, x + y, v_c)$  est équivalent à inclure  $(u, v, x, v_c)$  et  $(u, v, y, v_c)$ . Ceci est aussi utile pour les combinaisons linéaires, puisque l'hyperarc  $(u, v, x + \alpha y, v_c)$  est équivalent aux hyperarcs  $(u, v, x, v_c)$  et  $(u, v, y, v_c)$  avec le coefficient du second multiplié par  $\alpha$ . Cette séparation est également valide sur  $v_c$ .

**Définition 2.1.10.** Un hypergraphe de mouvement (ou carquois de mouvement) est un 3-tuple  $(V, A, f)$ , où  $(V, A)$  est un hypergraphe et  $f : A \rightarrow \mathbb{R}^+$  est un paramètre de mouvement.

**Définition 2.1.11.** Sur le digraphe ou hypergraphe  $G$  (de mouvement ou non), on définit les fonctions  $\underline{V}(G) : G \rightarrow V$  et  $\underline{A}(G) : G \rightarrow A$  qui donnent respectivement l'ensemble des nœuds et des hyperarcs de  $G$ . Ces fonctions seront utiles pour référer aux ensembles de nœuds et d'hyperarcs de deux structures distinctes par la même notation.

*Notation 2.1.12.* Pour représenter complètement un hyperarc  $e = (u, v, v_r, v_c)$  dans la figure d'un hypergraphe, on tracera d'abord une flèche pleine entre  $u$  et  $v$  (appelons-la  $u \rightarrow v$ ). Ensuite, on ajoutera une flèche pointillée à partir de chacun des nœuds dans la somme  $v_r$  vers la flèche  $u \rightarrow v$  et une flèche segmentée à partir de chacun des nœuds dans la somme  $v_c$  vers la flèche  $u \rightarrow v$ .

**Remarque 2.1.13.** L'utilisation du terme « carquois » s'inspire de la théorie des catégories, mais les deux définitions ne doivent pas être confondues. On utilise ce terme dans ce contexte parce que les hypergraphes sont représentés par une collection de flèches.

## 2.2. Carquois épidémiologiques

**Définition 2.2.1.** Soit un carquois  $G$  et un nœud  $v \in V(G)$ . On note les hyperarcs entrants de  $v$  avec  $A^-(v)$  et les hyperarcs sortants de  $v$  avec  $A^+(v)$ . Formellement,  $A^-(v) = \{(x, y, v_r, v_c) \in A(G) : y = v\}$  et  $A^+(v) = \{(x, y, v_r, v_c) \in A(G) : x = v\}$ .

**Définition 2.2.2.** Une source est un nœud  $u$  tel que  $A^-(u) = \emptyset$  (sans entrées) et un puits est un nœud  $v$  tel que  $A^+(v) = \emptyset$  (sans sorties).

**Définition 2.2.3.** Pour le reste du mémoire, les carquois auront toujours deux nœuds appelés nœuds vides, notés  $\text{Null}_n$  et  $\text{Null}_d$ , qui servent respectivement à créer les entrées et à accumuler les sorties de la population.  $\text{Null}_n$  est une source et  $\text{Null}_d$  est un puits.

*Notation 2.2.4.* Puisque  $\text{Null}_n$  et  $\text{Null}_d$  jouent essentiellement le même rôle et ne peuvent pas être placés aux mêmes endroits dans la notation des hyperarcs, on parlera simplement du nœud « Null » pour parler d'un de ces nœuds (lorsque cela ne porte pas à confusion).

**Définition 2.2.5.** On peut classifier les hyperarcs  $e \in A(G)$  dans quatre sous-ensembles qui partitionnent  $A(G)$ , batches, rates, births et contacts, définis de la manière suivante :

$$(u, v, v_r, v_c) \in \begin{cases} \text{batches}(G) & \text{si } v_r, v_c = \text{Null} \\ \text{rates}(G) & \text{si } v_r \neq \text{Null}, v_c = \text{Null}, u \neq \text{Null}_n \\ \text{births}(G) & \text{si } v_r \neq \text{Null}, v_c = \text{Null}, u = \text{Null}_n \\ \text{contacts}(G) & \text{si } v_r, v_c \neq \text{Null}. \end{cases} \quad (2.2.1)$$

C'est cette classification qui permettra une analyse automatique de la structure du graphe.

**Définition 2.2.6.** Les nœuds susceptibles d'un modèle  $G$  sont ceux qui sont influencés par les hyperarcs de contact, c'est-à-dire,  $S(A) = \bigcup_{e \in \text{contacts}} \{v_r(e)\}$ .

**Définition 2.2.7.** Les nœuds infectés (ou nœuds contagieux) d'un modèle  $G$  sont ceux qui causent les infections auprès des susceptibles, c'est-à-dire  $I(A) = \bigcup_{e \in \text{contacts}} \{v_c(e)\}$ .

**Définition 2.2.8.** Un carquois épidémiologique est un 3-tuple  $G = (V, A, f)$  où :

- 1)  $(V, A)$  est un carquois ; (Structure générale)
- 2)  $f : A(G) \rightarrow \mathcal{L}^+[t]$  est une fonction de mouvement ; (Paramètres du modèle)
- 3) Chaque nœud  $v$  est une fonction du temps  $v : \mathbb{R} \rightarrow \mathbb{R}$  ; (Taille des compartiments)
- 4)  $\exists \text{Null}_n \in V(G) : \text{Null}_n$  est une source ; (Source d'entrées)
- 5)  $\exists \text{Null}_d \in V(G) : \text{Null}_d$  est un puits ; (Accumulation de sorties)

- 6)  $\forall e \in A(G)$ , on a  $v_r \neq \text{Null}_d$  et  $v_c \neq \text{Null}_n$ ; (Placement unique de  $\text{Null}_n, \text{Null}_d$ )
- 7)  $\forall e \in A(G), u \neq \text{Null}_n \implies v_r = u$ ; (Proportionnalité des négatifs)
- 8)  $\forall e \in A(G), e \in \text{contacts} \implies v_r \neq \text{Null}_n$ ; (Contacts doivent avoir une cible)
- 9) Tous les chemins entre  $S(A)$  et  $I(A)$  doivent passer par un **contact**

où  $\mathcal{L}^+[t]$  est l'ensemble des fonctions positives, bornées et  $\mathcal{C}^1$  par morceaux sur  $\mathbb{R}$ .

*Notation 2.2.9.* On note l'ensemble des carquois épidémiologiques  $\mathcal{G}$ .

**Définition 2.2.10.** On écrira  $V(G)^*$  pour désigner les nœuds qui font partie de la population.

Pour les carquois non modifiés, on aura  $V(G)^* = V(G) - \{\text{Null}_n, \text{Null}_d\}$ .

**Remarque 2.2.11.** On aura  $\forall v \in V(G)$  et  $\forall w \notin V(G)^*, v+w = v$  et  $\sum w = w$ .  $\text{Null}_n + \text{Null}_d$  sera égal à  $\text{Null}_n$  ou  $\text{Null}_d$  selon le contexte (par exemple,  $\text{Null}_n$  dans  $v_r$  et  $\text{Null}_d$  dans  $v_c$ ). Ceci permettra d'éviter les sommes pour  $v_r$  et  $v_c$  qui contiennent  $\text{Null}_n$  ou  $\text{Null}_d$  sans être simplement égales à  $\text{Null}$ .

*Notation 2.2.12.* Les nœuds vides seront notés «  $\cdot$  » lorsqu'ils se trouvent à l'intérieur de tuples et il sera aussi standard de les omettre lorsqu'on représente les carquois. Les lignes pointillées seront aussi omises lorsque  $v_r = u$ . Voir exemple 2.2.13.

*Exemple 2.2.13.* Dans l'hypergraphe  $G$  de l'exemple 2.1.8, on peut désigner les nœuds  $c$  et  $d$  comme  $\text{Null}_n$  et  $\text{Null}_d$ . On peut aussi appeler le nœud  $a$  «  $S$  » et le nœud  $b$  «  $I$  », ce qui nous donne la figure 2.5. Les hyperarcs sont coloriées en accord avec la définition 2.2.5.

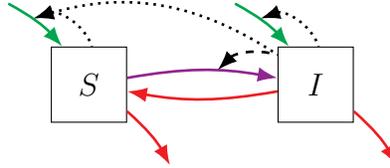


FIGURE 2.5 – Simplification de l'hypergraphe donné à l'exemple 2.1.8.

**Remarque 2.2.14.** La représentation de la figure 2.5 avec l'inclusion des nœuds  $\text{Null}$  et toutes les flèches est montrée à la figure 2.6. On voit maintenant de manière plus claire le rôle des nœuds vides et l'effet d'omettre les flèches grises.

**Remarque 2.2.15.** Pour une somme de nœuds  $\sigma = \sum v \in \Sigma(V)$ , on dira  $\sigma(t) = \sum v(t)$ . Ceci est naturel puisque les nœuds du graphe sont des fonctions et  $(f + g)(t) = f(t) + g(t)$ .

**Définition 2.2.16.** La population totale d'un carquois  $G$  au temps  $t$  est donnée par

$$N(G, t) = \sum_{v \in V(G)^*} v(t). \quad (2.2.2)$$

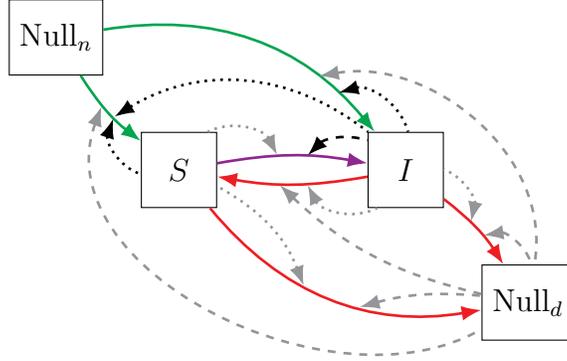


FIGURE 2.6 – Figure 2.5 avec l'inclusion des nœuds Null.

**Remarque 2.2.17.** Il est possible que certains nœuds soient à la fois susceptibles et infectés. Par exemple, on peut considérer un modèle à plusieurs maladies où, être infecté par une première maladie n'offre pas d'immunité pour la deuxième et on demeure susceptible.

## 2.3. Système d'équations différentielles

**Définition 2.3.1.** La variation d'un hyperarc  $e = (u, v, v_r, v_c)$  au temps  $t$  est donnée par

$$\Delta_e(t) = f(e)(t) \cdot r(e, t)c(e, t) \quad (2.3.1)$$

où l'influence de taux  $r$  est donnée par

$$r(e, t) = \begin{cases} v_r(t) & \text{si } v_r \neq \text{Null} \\ 1 & \text{si } v_r = \text{Null} \end{cases} \quad (2.3.2)$$

et l'influence de contact  $c$  est donnée par

$$c(e, t) = \begin{cases} v_c(t)/N(G, t) & \text{si } v_c \neq \text{Null} \\ 1 & \text{si } v_c = \text{Null}. \end{cases} \quad (2.3.3)$$

**Définition 2.3.2.** Pour tous les nœuds  $v$ , on définit l'équation différentielle de  $v$  comme

$$\frac{dv(t)}{dt} = \sum_{e^- \in A^-(v)} \Delta_{e^-}(t) - \sum_{e^+ \in A^+(v)} \Delta_{e^+}(t). \quad (2.3.4)$$

**Remarque 2.3.3.** Les deux nœuds vides ont aussi une valeur dictée par cette même équation différentielle. Leur valeur permettra de calculer la variation de la population totale, ainsi que le nombre d'entrées et de sorties cumulatives. Voir théorème 2.3.4 et corollaire 2.3.5.

**Théorème 2.3.4.** *Pour tout carquois épidémiologique  $G$ , on a*

$$N(G, t) - N(G, 0) = - \sum_{v \notin V(G)^*} (v(t) - v(0)). \quad (2.3.5)$$

DÉMONSTRATION. Pour tout nœud  $v \in V(G)$ , on a que tout hyperarc  $e$  sortant de  $v$ , il existe un unique nœud  $w$  de  $G$  tel que  $e$  est entrant pour  $w$ . Ceci implique que tous les mouvements du graphe s'annulent dans les équations différentielles, et donc  $\forall t \in \mathbb{R}$ ,

$$\begin{aligned} \sum_{v \in V(G)} \frac{dv(t)}{dt} = 0 &\implies \sum_{v \in V(G)^*} \frac{dv(t)}{dt} + \sum_{v \notin V(G)^*} \frac{dv(t)}{dt} = 0 \\ &\implies \frac{d}{dt} \sum_{v \in V(G)^*} v(t) = \frac{d}{dt} N(G, t) = - \sum_{v \notin V(G)^*} \frac{dv(t)}{dt}. \end{aligned}$$

En utilisant le théorème fondamental du calcul, on aura

$$\int_0^T \frac{d}{dt} N(G, t) dt = - \int_0^T \sum_{v \notin V(G)^*} \frac{dv(t)}{dt} dt = - \sum_{v \notin V(G)^*} \int_0^T \frac{dv(t)}{dt} dt$$

et donc

$$N(G, t) \Big|_0^T = - \sum_{v \notin V(G)^*} v(t) \Big|_0^T. \quad \blacksquare$$

*Corollaire 2.3.5.* Si les nœuds vides sont les seuls nœuds de  $V(G)$  qui ne sont pas dans  $V(G)^*$ , on aura

$$\Delta N = -(\Delta \text{Null}_n + \Delta \text{Null}_d). \quad (2.3.6)$$

DÉMONSTRATION. Cette démonstration est triviale. \blacksquare

**Remarque 2.3.6.** Puisque  $\text{Null}_n$  est une source et  $\text{Null}_d$  est un puits, on peut calculer le nombre cumulatif d'entrées dans le modèle avec  $\Delta \text{Null}_n$  et le nombre cumulatif de sorties avec  $-\Delta \text{Null}_d$ . Attention : le calcul des entrées sera différent avec l'ajout des nœuds pour calculer  $\mathcal{R}_t$ . Voir remarque 3.1.4.

*Exemple 2.3.7.* Soit le carquois de mouvement montré à la figure A.1. On déduit que  $A^-(\text{Null}_n) = A^+(\text{Null}_d) = \emptyset$ . On donne les valeurs suivantes pour  $f$  :

$$f(S, I, S, I) \equiv \beta, \quad f(I, S, I, \cdot) \equiv \gamma, \quad f(S, R, \cdot, \cdot) \equiv \nu, \quad f(R, S, R, \cdot) \equiv \delta.$$

On aura donc

$$\begin{aligned}
- \Delta(S, I, S, I)(t) &= \beta S \frac{I}{N}(t); & - \Delta(S, R, \cdot, \cdot)(t) &= \nu S(t); \\
- \Delta(I, S, I, \cdot)(t) &= \gamma I(t); & - \Delta(R, S, R, \cdot)(t) &= \delta R(t);
\end{aligned}$$

et finalement

$$\begin{aligned}
S'(t) &= \left( \gamma I + \delta R - \beta S \frac{I}{N} - \nu S \right)(t) \\
I'(t) &= \left( \beta S \frac{I}{N} - \gamma I \right)(t) \\
R'(t) &= (\nu S - \delta R)(t) \\
\left. \begin{array}{l} \text{Null}'_n(t) = 0 \\ \text{Null}'_d(t) = 0 \end{array} \right\} &\implies N'(t) = 0
\end{aligned} \tag{2.3.7}$$

ce qui nous redonne les équations standards qu'on aurait données aux nœuds du graphe.

## 2.4. Produits d'hypergraphes

Un avantage important de la structure proposée est la possibilité d'adapter le produit cartésien sur les graphes, tel que montré en [25], pour lui donner un comportement équivalent sur les hypergraphes. Par exemple, les carquois épidémiologique  $G$  et  $H$  peuvent être combinés à l'aide du produit cartésien  $\square$  pour donner le modèle  $G \square H$ , qui sera défini avec :

$$\begin{aligned}
V(G \square H) &= V(G) \times V(H), \\
A(G \square H) &= \bigcup_{w \in V(G)^*} \bigcup_{\substack{(u,v,v_r,v_c) \\ \in A(H)}} \left\{ \left( (w, u), (w, v), \sum_{x \in v_r} (w, x), \sum_{y \in V(G)^*} \sum_{x \in v_c} (y, x) \right) \right\} \\
&\quad \cup \bigcup_{w \in V(H)^*} \bigcup_{\substack{(u,v,v_r,v_c) \\ \in A(G)}} \left\{ \left( (u, w), (v, w), \sum_{x \in v_r} (x, w), \sum_{y \in V(H)^*} \sum_{x \in v_c} (x, y) \right) \right\}.
\end{aligned} \tag{2.4.1}$$

Le calcul de  $A(G \square H)$  peut possiblement être abrégé en prenant  $v + (\cdot, w) = v + (w, \cdot) = v$  pour tout nœud  $v$  dans  $V(G \square H)$ .

Pour conserver la structure de chaque modèle original, cette définition du produit cartésien est naturelle. En effet, pour chaque compartiment de la population du premier modèle, les hyperarcs du second modèle doivent être présents [25]. De plus, il faut que la proportionnalité soit conservée, et donc on pose  $(w, v_r)$  comme influence de taux, qu'il faut séparer en

somme pour respecter la définition de  $\Sigma(V)$ . Finalement, pour que l'influence de contact ne soit pas modifiée, il faut faire la somme de  $(w, v_c)$  sur chaque nœud  $w$ . Ceci assure que tous les individus de la première structure affecteront les susceptibles de la deuxième.

La dynamique du modèle résultant ne sera pas la même que celle des deux modèles initiaux, lorsqu'on compresse le modèle sur chaque dimension. Par exemple, le modèle de l'exemple 2.4.1 voit apparaître des mortalités pour  $S$  et  $R$  qui ne sont simplement pas présentes dans le modèle  $SIR$  considéré.

*Exemple 2.4.1.* Soient  $G$  et  $H$  deux carquois épidémiologiques définis par

$$\begin{aligned}
 V(G) &= \{A_0, A_1\}, \\
 A(G) &= \{(\cdot, A_0, A_0, \cdot), (A_0, \cdot, A_0, \cdot), (A_0, A_1, A_0, \cdot), (A_1, \cdot, A_1, \cdot)\}, \\
 V(H) &= \{S, I, R\}, \\
 A(H) &= \{(S, I, S, I), (I, R, I, \cdot), (I, \cdot, I, \cdot)\}.
 \end{aligned}
 \tag{2.4.2}$$

Le graphe  $G$  donne une structure d'âge simple, qui prend en compte la démographie de la population. Le graphe  $H$  est un modèle  $SIR$  où la seule démographie est celle causée par la maladie. Individuellement, ces graphes donnent un portrait très peu réaliste d'une population en situation d'épidémie. Cependant, ils peuvent être combinés pour donner le schéma de la figure 2.7, qui représente un modèle beaucoup plus robuste.

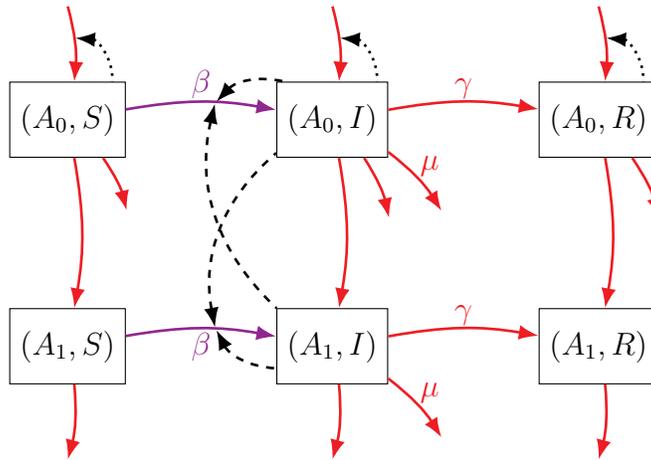


FIGURE 2.7 – Produit de deux hypergraphes  $G$  et  $H$ .

## 2.5. Structure informatique

De manière générale, on peut stocker l'information nécessaire pour un carquois  $G$  à l'intérieur d'un fichier « nom du modele.json », contenant l'information suivante :

```
1 {
2   "name": "nom du modele",
3   "compartments": {
4     "compartiment_1": condition_initiale_1,
5     ... autres compartiments
6   },
7   "flows": {
8     "categorie_1": [
9       {
10        "from": "compartiment_i",
11        "to": "compartiment_j",
12        "rate": "compartiment_k + ...",
13        "contact": "compartiment_l + ...",
14        "parameter": "fonction du temps"
15      },
16      ... autres arretes
17    ],
18    ... autres categories
19  }
20 }
```

Ceci nous permet de contenir toutes les informations nécessaires et de regrouper les hyperarcs en catégories (qui n'ont pas d'influence sur la structure du modèle). Celles-ci peuvent permettre une lecture facile du fichier pour les modèles de plus grande taille.

Par exemple, un modèle *SIR* avec deux compartiments contagieux en chaîne et un taux de contact cyclique d'une période de 30 jours (voir section 6.2.5 pour l'étude d'un modèle similaire) peut être stocké dans ce fichier « SIR\_sin\_2.json » :

```
1 {
2   "name": "SIR_sin_2",
3   "compartments": {
4     "S": 10000000,
```

```

5     "I1": 1,
6     "I2": 0
7     "R": 0,
8     "Null_n": 0,
9     "Null_m": 0,
10  },
11  "flows": {
12    "infections": [
13      {
14        "from": "S",
15        "to": "I1",
16        "rate": "S",
17        "contact": "I1+I2",
18        "parameter": "0.5 + 0.4*sin(2*pi*t/30)"
19      }
20    ],
21    "migrations": [
22      {
23        "from": "I1",
24        "to": "I2",
25        "rate": "I1",
26        "contact": "Null_m",
27        "parameter": "0.4"
28      }
29    ],
30    "recoveries": [
31      {
32        "from": "I2",
33        "to": "R",
34        "rate": "I2",
35        "contact": "Null_m",
36        "parameter": "0.4"
37      }
38    ]
39  }
40 }

```

Il est important de noter les limitations suivantes au niveau de l'implémentation des fichiers `json` :

- 1) Les compartiments ne peuvent pas contenir le caractère «  $\wedge$  ». Celui-ci est réservé pour désigner la couche 0 et la couche 1 pour les modèles modifiés ;
- 2) Les compartiments ne peuvent pas commencer par « `Null` » ou « `Rt` » ;
- 3) Le nom du fichier `json` et le paramètre "`name`" dans le fichier doivent être le même. Ceci est utile pour créer le fichier du modèle modifié (ce qui n'est pas nécessaire pour calculer les nombres de reproduction) ;
- 4) Les combinaisons linéaires de nœuds doivent absolument avoir des coefficients égaux à 0 ou 1, pour la lecture des sommes à partir des *strings*. La séparation de combinaisons linéaires en plusieurs hyperarcs (voir remarque 2.1.9) peut rendre difficile la modification manuelle des modèles ;
- 5) L'écriture de fonctions complexes pour le paramètre des hyperarcs n'est pas naturelle, puisqu'il faut tout écrire en une seule fonction contenue sur une ligne ;
- 6) Tous les termes dans les équations différentielles sont restreints à la forme  $f(t)uv/N$ . Les fonctions de capacité maximale décrites au chapitre 5 permettent l'ajout de termes plus complexes, mais la notation demeure limitative.

Malgré ce dernier point, la simplicité des termes d'équations différentielles assure les points suivants :

- 1) Il suffit de préciser quels nœuds causent les infections et créent les entrées du modèle pour complètement définir les termes d'ODEs ;
- 2) Il sera plus simple de créer une interface graphique interactive, où on peut ajouter des nœuds et des hyperarcs pour créer des modèles et générer les fichiers plus facilement ;
- 3) L'utilisation de  $v_r$  et  $v_c$  permet une caractérisation des hyperarcs en catégories, ce qui permettra une analyse simplifiée du modèle au chapitre 3.

# Chapitre 3

---

## Calcul explicite de $\mathcal{R}_0$ et $\mathcal{R}_t$

Environ 70 ans après l'introduction des nombres de reproduction par Alfred Lotka [1], Diekmann et al. [6] introduisent en 1990 la méthode de la prochaine génération (next-generation method) qui permet d'obtenir une forme analytique pour le nombre de reproduction de base d'un modèle à compartiments. En 2002, van den Driessche et Watmough [23] reformulent cette technique pour la généraliser et en 2005, Heffernan et al. [13] réécrivent cette méthode de manière plus pédagogique [16]. Diekmann et al. [7] précisent en 2009 comment bien appliquer la méthode pour certains modèles, pour « dissiper la confusion »<sup>1</sup> associée à celle-ci. Suite à son introduction, la méthode de la prochaine génération est devenue la méthode de référence pour calculer le nombre de reproduction de base.

Garnett (2002 et 2005) [9, 10] présente finalement l'égalité  $\mathcal{R}_t = \mathcal{R}_0 x$ , où  $x = S/N$ , qui est souvent utilisée pour calculer le nombre de reproduction effectif [12, 17]. Cependant, cette formule n'est pas mentionnée dans certains articles plus récents qui discutent de  $\mathcal{R}_t$  pour la COVID-19 [15, 19]. L'INSPQ présente sur son site une courte section qui explique le calcul des nombres de reproduction pour la pandémie de 2020 [20].

Dans ce chapitre, le but sera d'utiliser la structure de graphes définie au chapitre 2 pour établir un calcul des nombres de reproduction  $\mathcal{R}_0$  et  $\mathcal{R}_t$ . Ce nouveau calcul sera plus facile à vérifier que celui basé sur la matrice de prochaine génération, puisque la quantité sera calculée explicitement par la structure du modèle. De plus, il sera plus facile à adapter à une variété de modèles, puisqu'il ne repose pas sur une analyse mathématique propre à la structure de chacun.

---

1. Formulation originale : *We clear up confusion that has been around in the literature concerning the construction of this matrix, specifically for the most frequently used so-called compartmental models.*

Pour calculer les nombres de reproduction associés à un modèle de manière explicite, il sera important de mettre en évidence certains individus. On isolera de manière virtuelle (c'est-à-dire au niveau des équations différentielles, sans réellement influencer les interactions entre les individus) les nouveaux infectés créés au temps  $t$  dans un ensemble de compartiments copiés de ceux du modèle original. Il faudra ensuite comptabiliser le nombre total de cas secondaires générés par ces individus infectés originaux.

À la fin de son infection, un individu aura infecté (dans un modèle déterministe)  $\mathcal{R}_t$  individus. L'individu fera ensuite un retour dans un compartiment non contagieux et ne sera plus à l'étude pour le calcul des cas secondaires à ce moment. Pour bien définir la fin d'une infection, on utilisera le terme « chaîne d'infection » pour désigner les compartiments atteignables entre le début d'une infection et le moment où il est impossible de redevenir contagieux sans se faire infecter de nouveau.

Cette explication donne place à trois étapes nécessaires pour le calcul de  $\mathcal{R}_t$  :

- 1) Déterminer quels compartiments font partie de la chaîne d'infection de la maladie ;
- 2) Écrire un algorithme pour une modification du modèle qui duplique la chaîne d'infection et les hyperarcs qui s'y rattachent sans modifier la dynamique du modèle ;
- 3) Inclure dans le modèle un compteur pour accumuler le nombre d'infectés produits spécifiquement par les individus isolés, qui ne modifie pas la dynamique du modèle.

La chaîne d'infection peut être déterminée de la manière suivante : tous les compartiments à la fin d'une chaîne (en considérant un tri topologique du carquois  $G$ ) seront compris dans le nœud  $v_c$  d'un hyperarc de type **contacts**. Si ce n'était pas le cas, ces compartiments ne causeraient pas d'infection et on pourrait simplement les exclure de la chaîne puisqu'ils sont à la fin de celle-ci et n'ont pas d'importance pour la transmission. De plus, les compartiments au début d'une chaîne sont ceux qui viennent d'être infectés directement par un certain compartiment de la chaîne, et donc chacun est le nœud d'arrivée d'un **contact**.

Pour chaque hyperarc de type **contacts**, on sait donc que l'ensemble des nœuds compris entre  $v$  et  $v_c$  feront partie de la chaîne d'infection. On peut donc appliquer un algorithme de recherche pour trouver tous les chemins entre  $v$  et  $x$  qui sont contenus dans  $A(G) \setminus \text{contact}(G)$ , pour tous les  $x \in v_c$ . En prenant l'union de ces nœuds pour l'ensemble des **contacts**, on calculera notre chaîne d'infection. Par exemple, pour le modèle de la figure 3.1, on a seulement l'hyperarc de contact  $(S, E, S, I)$  qui nous donne la chaîne d'infection  $\{E, I\}$ .



FIGURE 3.1 – Modèle  $SEIR$  simple.

Quand la chaîne d'infection a été déterminée, on peut ajouter une version distincte de tous ses nœuds pour créer une couche isolée. Pour ce qui est de la création des nouveaux hyperarcs, on donne à la définition 3.1.3 une méthode pour dupliquer les hyperarcs nécessaires en maintenant la dynamique du carquois. On montre à la section 4.3 le bon fonctionnement de cette méthode.

Dans cette même définition, on ajoute aussi les compteurs pour les infections générées par la couche isolée. Plutôt que de compiler ces infections totales après avoir généré les courbes solution (à l'aide d'une intégrale bien posée), il sera utile d'utiliser des nœuds additionnels, notés  $\mathcal{R}_t(e)$ , qui accumulent spécifiquement le nombre à calculer. En effet, supposons que le nombre d'infectés générés par un hyperarc au temps  $t$  soit  $f(t)$ , on peut calculer  $\int_a^b f(t) dt$  à l'aide d'un nœud  $v$  initialisé avec une valeur  $v(a) = 0$  et dont l'équation différentielle est  $v'(t) = f(t)$ . On peut finalement extraire  $v(b)$  pour connaître la valeur de l'intégrale.

Soient  $e = (u, v, v_r, v_c) \in \text{contacts}(G)$  et  $C$  la chaîne d'infection de  $G$ . On aura alors  $\{v_r, v_c\} \subseteq C$ , et donc il existe des nœuds  $v'_r, v'_c$  qui sont des copies des nœuds originaux (et les nœuds originaux sont toujours dans le graphe). Pour observer l'impact de la couche isolée sur les infections, il faut donc calculer les infections générées par  $v'_c$  sur  $v_r$ , c'est-à-dire

$$\int_a^\infty f(e) \cdot v_r \frac{v'_c}{N} dt. \quad (3.0.1)$$

### 3.1. Algorithme d'ajout aux hypergraphes

Soit un hypergraphe de mouvement  $G = (V, A, f)$ . On souhaite créer  $G' = (V', A', f') = \text{mod}(G)$ , un nouveau hypergraphe de mouvement qui respecte la même dynamique que  $G$  et qui permet de séparer un certain nombre d'infectés particuliers du reste de la population.

**Définition 3.1.1.** Une sous-chaîne infectieuse de  $G$  entre deux nœuds, notée  $E_{u \rightarrow v}$ , est défini comme l'union des nœuds contenus dans tous les chemins de  $u$  à  $v$  qui ne passent pas par des hyperarcs dans  $\text{contacts}(G)$ . Soient  $u_0 = u$  et  $u_n = v$ . Pour tout chemin  $p = u_0 u_1 \cdots u_n$  tel que  $(u_i, u_{i+1}) \notin \text{contacts}(G)$ , on aura  $u_i \in E_{u \rightarrow v}$ . Voir algorithmes 1 à 4 de l'annexe B.

**Définition 3.1.2.** On note  $\mathcal{U}$  le sous-ensemble de  $\mathcal{G}$  des modèles non-modifiés.

**Définition 3.1.3.** Soit  $G = (V, A, f)$ . On définit  $\text{mod} : \mathcal{U} \rightarrow \mathcal{G}$ , où  $\text{mod}(G) = (V', A', f')$ , en donnant d'abord  $V'$  et  $(V')^*$  :

$$\begin{aligned} \omega &= \{w^1 : w \in V(G)^*\} \quad (\text{originaux}) & \iota &= \bigcup_{e \in \text{contacts}(G)} \{w^0 : w \in E_{v(e) \rightarrow v_c(e)}\} \quad (\text{chaîne}) \\ \nu &= \{\text{Null}_n, \text{Null}_d\} \quad (\text{nœuds Null}) & \kappa &= \{\mathcal{R}_t(e) : e \in \text{contacts}(G)\} \quad (\text{compteurs}) \\ V' &= \omega \cup \iota \cup \nu \cup \kappa \quad \text{et} & (V')^* &= \omega \cup \iota = V' - (\nu \cup \kappa). \end{aligned} \quad (3.1.1)$$

On peut vérifier que  $\omega$ ,  $\iota$ ,  $\nu$  et  $\kappa$  sont des ensembles disjoints. On donne ensuite les définitions suivantes pour séparer les nœuds dans le nouveau graphe :

$$i_v = \min\{j : v^j \in V'\} \quad \text{et} \quad \text{split}(v_r) = \sum_{w \in v_r} \sum_{i=1}^{i_w} w^i, \quad (3.1.2)$$

puis la définition de  $A'$  :

$$\begin{aligned} A' &= \{e' = (u^1, v^1, v_r^1, \cdot) \in \mathcal{A}(V') \mid e = (u, v, v_r, \cdot) \in \text{rates}(G)\} \\ &\cup \{e' = (u^0, v^{i_v}, v_r^0, \cdot) \in \mathcal{A}(V') \mid e = (u, v, v_r, \cdot) \in \text{rates}(G) : u^0 \in V'\} \\ &\cup \{e' = (u^1, v^1, v_r^1, \text{split}(v_c)) \in \mathcal{A}(V') \mid e = (u, v, v_r, v_c) \in \text{contacts}(G)\} \\ &\cup \{e' = (u^0, v^0, v_r^0, \text{split}(v_c)) \in \mathcal{A}(V') \mid e = (u, v, v_r, v_c) \in \text{contacts}(G) : u^0 \in V'\} \\ &\cup \{e' = (\cdot, v^1, \text{split}(v_r), \cdot) \in \mathcal{A}(V') \mid e = (\cdot, v, v_r, \cdot) \in \text{births}(G)\} \\ &\cup \{e' = (\cdot, v^1, \cdot, \cdot) \in \mathcal{A}(V') \mid e = (\cdot, v, \cdot, \cdot) \in \text{batches}(G)\} \\ &\cup \{e' = (\cdot, \mathcal{R}_t(e), \text{split}(v_r), v_c^0) \in \mathcal{A}(V') \mid e = (u, v, v_r, v_c) \in \text{contacts}(G)\} \end{aligned} \quad (3.1.3)$$

où  $\mathcal{A}(V')$  est donnée à l'équation 2.1.2. On définit  $f'(e') = f(e)$  pour chaque  $e' \in A'$ .

**Remarque 3.1.4.** On exclut les nœuds  $\mathcal{R}_t(\dots)$  de  $(V')^*$  puisqu'ils ne représentent pas des individus, et ainsi le théorème 2.3.4 reste valide. De plus, on a toujours que  $\Delta\text{Null}_d = \#\{\text{mortalités}\}$  et on peut calculer le nombre d'entrées avec

$$\#\{\text{entrées}\} = - \sum_{v \notin V(G)^*} \Delta v + \Delta\text{Null}_d.$$

*Exemple 3.1.5.* Les figures 3.2 et 3.3 montrent un exemple de modification en utilisant l'algorithme de la définition 3.1.3.

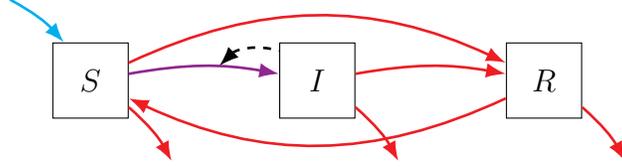


FIGURE 3.2 – Modèle *SIRS* avec démographie et vaccination.

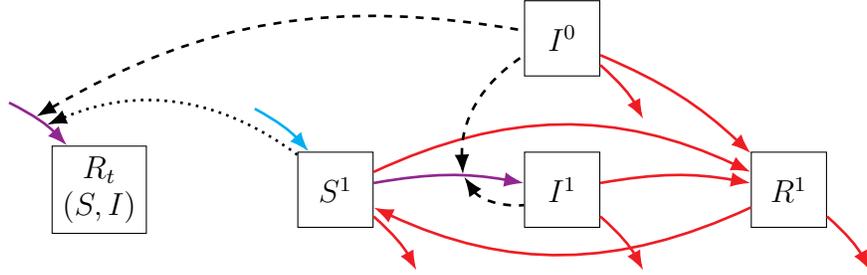


FIGURE 3.3 – Modification du modèle *SIRS* de la figure 3.2.

**Remarque 3.1.6.** On rappelle que la couche 0 contient un ensemble d'individus que l'on souhaite isoler du reste de la population. Il faut cependant s'assurer que cette séparation n'affecte pas la dynamique du modèle, sans quoi on n'aura aucune assurance que les nombres de reproduction calculés seront ceux associés à l'épidémie qui nous intéresse. On définit maintenant ce que « maintenir la dynamique » signifie.

**Définition 3.1.7.** Soit  $G$  un carquois et  $G' = \text{mod}(G)$ . On dira que  $G$  et  $G'$  ont la même dynamique si tous les nœuds  $v$  de la population  $V(G)^*$  respectent

$$\forall t \geq 0, v(t) = \text{split}(v)(t) \quad (3.1.4)$$

c'est-à-dire si les tailles des compartiments du carquois non-modifié sont « bien répartis » entre les couches 0 et 1 du carquois modifié.

*Exemple 3.1.8.* Il est possible de vérifier cette propriété à la main sur le carquois de l'exemple 3.1.5. Soient  $S(0) = S^1(0)$ ,  $I(0) = I^0(0) = 1$ ,  $R(0) = I^1(0) = R^1(0) = 0$ . On peut simplement vérifier que la dérivée de chaque compartiment  $v$  est la même que  $\text{split}(v)$ .

– Pour  $S$  :

$$\frac{dS(t)}{dt} = \left( \Lambda + \delta R - \nu S - \mu S - \beta S \frac{I}{S + I + R} \right)(t)$$

et  $\frac{d\text{split}(S)(t)}{dt} = \frac{dS^1(t)}{dt} = \left( \Lambda + \delta R^1 - \nu S^1 - \mu S^1 - \beta S^1 \frac{I^0 + I^1}{S^1 + (I^0 + I^1) + R^1} \right)(t).$

– Pour  $I$  :

$$\begin{aligned}\frac{dI(t)}{dt} &= \left( \beta S \frac{I}{S + I + R} - \gamma I - \mu I \right)(t) \\ \text{et } \frac{d \text{split}(I)(t)}{dt} &= \frac{dI^0(t)}{dt} + \frac{dI^1(t)}{dt} \\ &= \left( -\gamma I^0 - \mu I^0 \right)(t) + \left( \beta S^1 \frac{I^0 + I^1}{S^1 + (I^0 + I^1) + R^1} - \gamma I^1 - \mu I^1 \right)(t) \\ &= \left( \beta S^1 \frac{I^0 + I^1}{S^1 + (I^0 + I^1) + R^1} - \gamma (I^0 + I^1) - \mu (I^0 + I^1) \right)(t).\end{aligned}$$

– Pour  $R$  :

$$\begin{aligned}\frac{dR(t)}{dt} &= (\gamma I + \nu S - \delta R - \mu R)(t) \\ \text{et } \frac{d \text{split}(R)(t)}{dt} &= \frac{dR^1(t)}{dt} = (\gamma I^0 + \gamma I^1 + \nu S^1 - \delta R^1 - \mu R^1)(t) \\ &= (\gamma (I^0 + I^1) + \nu S^1 - \delta R^1 - \mu R^1)(t)\end{aligned}$$

ce qui concorde avec le résultat voulu. Une preuve générale sera donnée à la section 4.3.

**Définition 3.1.9.** La base d'un nœud d'un carquois modifié est définie comme  $b(v^i) = v$ , c.-à-d. le nœud dans le modèle original qui a amené la création de  $v^i$  dans le carquois modifié. On aura  $b(\text{Null}_n) = \text{Null}_n$  et de même pour  $\text{Null}_d$ . Les nœuds  $\mathcal{R}_t(\dots)$  n'ont pas de base et par convention on dira  $b(\mathcal{R}_t(e)) = \text{Null}$ .

**Remarque 3.1.10.** En pratique (par exemple, dans l'implémentation en Python), la base d'un nœud peut être trouvée simplement en lui enlevant son indice. Il n'est pas nécessaire de connaître le multidigraphe de base pour appliquer la fonction. Cependant, il sera souvent connu, puisqu'on l'utilise pour construire le multidigraphe modifié.

**Définition 3.1.11.** La couche de  $v' \in V(G')$  est définie comme l'indice supérieur de  $v'$ . Elle sera notée  $c(v) \in \mathbb{Z}$  et on écrira parfois  $v \in c(a)$  si  $c(v) = a$  (par abus de notation). On dira que  $\text{Null} \in c(-1)$  et  $\forall i, \text{Null}^i = \text{Null}$ , et de même pour les nœuds  $\mathcal{R}_t(\dots)$ .

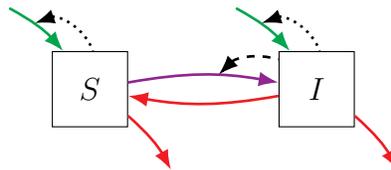


FIGURE 3.4 – Modèle  $SIS$  avec démographie.

*Exemple 3.1.12.* Soit le modèle de la figure 3.4. En utilisant la définition 3.1.3, on trouve les nouveaux hyperarcs (tableau 3.1) et on construit le carquois modifié de la figure 3.5.

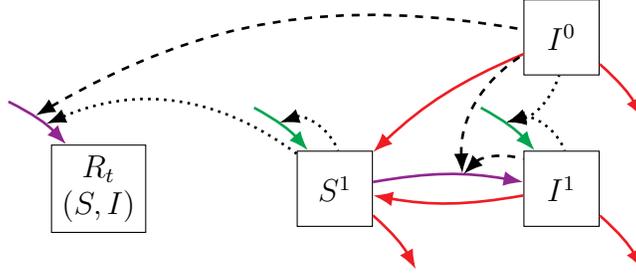


FIGURE 3.5 – Modification d'un modèle  $SIS$  avec démographie (figure 3.4).

$(e, f(e))$	$(e', f'(e'))$
$(S, I, S, I)$	$(S^1, I^1, S^1, I^0 + I^1)$ $(\cdot, \mathcal{R}_t(S, I), S^1, I^0)$
$(I, S, I, \cdot)$	$(I^0, S^1, I^0, \cdot)$ $(I^1, S^1, I^1, \cdot)$
$(\cdot, S, S, \cdot)$	$(\cdot, S^1, S^1, \cdot)$
$(\cdot, I, I, \cdot)$	$(\cdot, I^1, I^0 + I^1, \cdot)$
$(S, \cdot, S, \cdot)$	$(S^1, \cdot, S^1, \cdot)$
$(I, \cdot, I, \cdot)$	$(I^0, \cdot, I^0, \cdot)$ $(I^1, \cdot, I^1, \cdot)$

TABLEAU 3.1 – Nouveaux hyperarcs pour le modèle de la figure 3.4.

## 3.2. Conditions initiales du modèle modifié

On s'intéresse maintenant à l'initialisation des nœuds du carquois modifié. Supposons qu'on veuille calculer le  $\mathcal{R}_\tau$  associé au modèle généré par le graphe  $G$  pour un certain  $\tau \geq 0$ . Il faudra d'abord résoudre le système d'équations différentielles associé à  $G$ , c'est-à-dire

$$\left\{ \frac{dv(t)}{dt} = \sum_{e^- \in A^-(v)} \Delta_{e^-}(t) - \sum_{e^+ \in A^+(v)} \Delta_{e^+}(t) \mid v \in V(G) \right\} \quad (3.2.1)$$

sur  $0 \leq t \leq T$  avec  $T > \tau$ , avec un ensemble de conditions initiales. Pour calculer  $\mathcal{R}_\tau$ , on voudra ensuite extraire de cette solution les valeurs de tous les nœuds au temps  $\tau$ . Ensuite, on souhaite placer ces individus dans le modèle modifié en plaçant une partie des infectés dans la couche 0 pour ainsi tracer ces cas index.

On aura d'abord pour chaque hyperarc  $e \in A(G)$

$$\Delta_{\text{inf}}(e)(\tau) = \begin{cases} \Delta_e(\tau) & \text{si } e \in \text{contacts}(G) \\ 0 & \text{si } e \notin \text{contacts}(G) \end{cases} \quad (3.2.2)$$

qui représente le nombre de nouveaux infectés créés par  $e$  au temps  $\tau$ . Pour chaque nœud  $v \in V(G)^*$ , on donne l'équivalent

$$\delta_{\text{inf}}(v, \tau) = \sum_{e: v \in v_c(e)} \Delta_e(\tau). \quad (3.2.3)$$

On souhaite maintenant « intercepter » les nouveaux infectés et les placer dans la couche 0 plutôt que dans la couche 1. On définit donc les conditions initiales

$$v(\tau) = \begin{cases} \delta_{\text{inf}}(b(v), \tau) & \text{si } v \in c(0) \\ b(v)(\tau) - \delta_{\text{inf}}(b(v), \tau) & \text{si } v \in c(1) \\ b(v)(\tau) & \text{si } v = \text{Null} \\ 0 & \text{pour les autres nœuds} \end{cases} \quad (3.2.4)$$

et on résout le système 3.2.1, où les valeurs des paramètres  $f(e)(t)$  sont fixées à leur valeur lorsque  $t = \tau$ . On peut modifier l'équation 2.3.1 pour écrire

$$\Delta_e(t, t_0) = f(e)(t_0) \cdot r(e, t) c(e, t)$$

et ainsi réécrire le système d'EDOs 3.2.1 comme ceci :

$$\left\{ \frac{dv(t)}{dt} = \sum_{e^- \in A^-(v)} \Delta_{e^-}(t, \tau) - \sum_{e^+ \in A^+(v)} \Delta_{e^+}(t, \tau) \mid v \in V(G') \right\} \quad (3.2.5)$$

qu'on simule pour  $\tau \leq t \leq T$  avec  $T \gg \tau$ , avec les conditions initiales données en 3.2.4. On note aussi le nombre de nouveaux infectés créés au temps  $\tau$  par

$$\nu(\tau) = \sum_{w \in V(G)^*} \delta_{\text{inf}}(w, \tau). \quad (3.2.6)$$

En espérant que  $\mathcal{R}_t(e)(\infty) \approx \mathcal{R}_t(e)(T)$ , on s'intéresse alors aux valeurs

$$\mathcal{R}_\tau = \left\{ \frac{\mathcal{R}_t(e)(T)}{\nu(\tau)} \mid e \in \text{contacts} \right\}. \quad (3.2.7)$$

Cette convergence des nœuds  $\mathcal{R}_t(e)$  est étudiée à la section 4.5.

On répète ensuite ce processus pour toutes les valeurs de  $\tau$  pour lesquelles on souhaite connaître  $\mathcal{R}_\tau$ . En pratique, il sera suffisant de prendre tous les  $\tau \in [0, T_0] \cap \mathbb{Z}$ . Il est à noter que la première simulation n'a pas besoin d'être faite pour chacun des temps  $\tau$ . Il faut simplement recommencer le processus à partir de l'initialisation du modèle modifié.

### 3.3. Sur la normalisation de $\mathcal{R}_t$

Comme l'indique l'équation 3.2.7, le calcul de  $\mathcal{R}_\tau$  se fait avec

$$\frac{\mathcal{R}_t(e)(\infty)}{\nu(\tau)}. \quad (3.3.1)$$

Cette division vient du fait que les nœuds  $\mathcal{R}_t(e)$  accumulent le nombre d'infectés total générés par tous les individus de la couche 0. Cependant, la définition des nombres de reproduction est basée sur une seule personne, et il faut donc calculer combien de ces cas sont générés par un individu en moyenne, ce qui équivaut à diviser par  $\nu(\tau)$ .

Une autre solution, utilisée dans la première implémentation du code, est de placer exactement une personne dans la couche 0. Cela équivaut à normaliser  $\delta_{\text{inf}}$  par  $1/\nu(\tau)$  et donc la normalisation finale de  $\mathcal{R}_t$  n'est pas nécessaire. Cependant, cela crée un problème dans certains cas extrêmes où le nombre d'infectés total est inférieur à 1. En effet, si

$$\exists \tau \geq 0 : \sum_{e \in \text{contacts}(G)} v_c(e)(\tau) < 1, \quad (3.3.2)$$

on aura alors une valeur initiale négative dans la couche 1.

Pour la méthode proposée dans la section 3.2, on voit que ce problème disparaît. En effet, les contacts pour l'hyperarc  $e = (u, v, v_r, v_c)$  sont toujours créés par un terme de la forme

$$\beta(\tau) \cdot \frac{v_r(\tau)v_c(\tau)}{N(\tau)} = \frac{\beta(\tau)v_r(\tau)}{N(\tau)}v_c(\tau). \quad (3.3.3)$$

Donc,  $\forall e \in \text{contacts}(G)$ ,  $\forall v \in v_r(e)$ , tant que

$$\sum_{e: v \in v_c(e)} \frac{\beta(\tau)v_r(\tau)}{N(\tau)} < 1, \quad (3.3.4)$$

on aura une situation sans problème *a priori*. En particulier, si tous les  $\beta(\tau)$  sont  $< 1$ , ceci sera toujours vérifié. De plus, cela n'amène aucun changement au niveau de la dynamique, ce qui sera montré à la section 4.4.

## 3.4. Complexité des algorithmes

### 3.4.1. Méthode par simulation

Le nombre d'opérations pour générer une courbe de  $\mathcal{R}_t$  dépend grandement de la structure du modèle ainsi que de la complexité des coefficients donnés. Pour observer l'impact réel de ces quantités, rappelons le processus suivi :

- 1) Simuler le système d'équations différentielles pour le modèle de base.
- 2) Créer le modèle modifié.
- 3) Pour chaque moment  $t$ , initialiser le modèle modifié.
- 4) Pour chaque moment  $t$ , simuler le système d'EDOs modifié.

Puisque l'algorithme de `odeint` utilise un processus de retour arrière pour corriger les erreurs qui peuvent être introduites par une précision trop grossière, il est impossible de calculer la complexité exacte de l'algorithme sur une période donnée (tel qu'expliqué par Chris Rackauckas [21]). Cependant, si on se limite à certaines variables d'intérêt, on peut émettre l'hypothèse suivante :

**Conjecture 3.4.1.** *La complexité  $C$  de l'algorithme respecte*

$$C \in \mathcal{O}(n \cdot T \cdot |G|) \tag{3.4.1}$$

où  $n$  est le nombre de points pour lesquels on souhaite calculer  $\mathcal{R}_t$ ,  $T = t_f - t_i$  est l'intervalle de temps de simulation utilisé, et  $|G| = |V(G)| + |A(G)|$  est la taille totale du graphe.

Cette hypothèse vient intuitivement du fait que pour chaque temps  $t$  où on souhaite calculer  $\mathcal{R}_t$ , il faudra simuler un ensemble d'équations différentielles sur un temps  $T$ . Il est naturel que la complexité suive linéairement le temps de simulation, puisque la complexité de simulation pour un temps donné et la probabilité de devoir faire un retour arrière pour corriger une erreur possible ne dépend pas des temps précédents. À chaque mise à jour des nœuds, il faut calculer  $|A(G)|$  termes et  $|V(G)|$  nouvelles valeurs.

Il est à noter que cette hypothèse est simplement utilisée ici à titre de vérification pour ne pas utiliser un algorithme qui grandit trop rapidement en complexité en fonction du modèle étudié ou du temps de simulation. Pour les besoins de cette recherche, un algorithme de complexité linéaire est raisonnable.

### 3.4.2. Matrice de prochaine génération

Pour discuter de la complexité de la méthode de la matrice de prochaine génération, on commence par présenter sommairement cette méthode. Cette explication est tirée directement de Diekmann et al. (2009) [7].

Premièrement, on pose  $x$  comme le vecteur contenant tous les compartiments infectés, puis on réécrit la partie du système qui agit sur ces compartiments sous la forme linéarisée  $x' = (T + \Sigma)x$ . Ici, l'ensemble des termes qui créent de nouvelles infections sont placés dans la matrice  $T$ , et les termes de mouvement entre les compartiments infectés sont placés dans la matrice  $\Sigma$ . Lorsque ces deux matrices sont construites, il faut finalement calculer le produit  $K_L = -T\Sigma^{-1}$  et calculer la valeur propre dominante de cette matrice résultante.

L'analyse de complexité pour la matrice de prochaine génération a peu d'importance en pratique, puisque la construction des matrices est une étape importante de la méthode qui n'est pas automatisée à l'aide d'un algorithme. La rapidité dépend donc entièrement de l'étude du modèle effectuée par des mathématiciens, qui est habituellement un processus assez ardu lorsque les modèles sont de grande taille.

En comparaison avec notre méthode, qui ne requiert aucune étude de structure, la rapidité de cette méthode est nettement inférieure. En effet, l'ensemble des résultats produits dans ce mémoire, qui totalisent la création de courbes  $\mathcal{R}_t$  pour 8 modèles différents, peuvent être produits en moins de 2 minutes. En particulier, la courbe  $\mathcal{R}_t$  la plus demandante est créée en 15 secondes directement à partir du fichier `json` et la moins demandante est créée en moins de 3 secondes, ce qui ne se compare pas avec le temps d'analyse humaine d'un modèle.

L'avantage d'une méthode automatique est donc remarquable, et on verra au chapitre 6 que celle-ci offre également d'autres avantages dans son application.



# Chapitre 4

---

## Propositions et preuves

### 4.1. Unicité des solutions

Pour montrer le théorème 4.3.1, il sera important de montrer que la solution du système décrit en 2.3.4 est unique. En fait, il sera aussi important de connaître son unicité par principe, puisqu'on aimerait obtenir une seule modélisation possible de l'épidémie peu importe les conditions initiales des compartiments (nous travaillons après tout avec des modèles déterministes). Pour montrer ceci, on utilise le théorème de Cauchy-Lipschitz [14].

**Théorème 4.1.1.** *Soit  $\Omega = \mathbb{R} \times (\mathbb{R}^{>0})^n$  un ouvert, où  $n = |V(G)|$ . Soit  $f : \Omega \rightarrow \mathbb{R}^n$ , telle que  $f(t, X)$  est  $\mathcal{C}^1$  par morceaux sur  $t$  et  $\mathcal{C}^1$  sur  $X$ . Soient*

$$D = \{d_k \in \mathbb{R} : 1 \leq k \leq M \leq \infty\} \quad (4.1.1)$$

les points de discontinuité de  $f$  en  $t$ . Étant donnée les conditions initiales

$$I = \{(t_i, x_i) \in \Omega : 0 \leq i \leq M \leq \infty, d_i < t_i < d_{i+1}\}, \quad (4.1.2)$$

$X' = f(t, X)$  admet une solution unique sur  $(\mathbb{R} \setminus D) \times (\mathbb{R}^{>0})^n$ .

DÉMONSTRATION. On sait que la dérivée de chaque composante de  $X$  sera sous la forme

$$v'_i = \sum_{e \in A^-(v_i)} f(e)r(e)c(e) - v_i \sum_{e \in A^+(v_i)} f(e)c(e), \quad (4.1.3)$$

où  $f(e)$  ne dépend pas des valeurs  $r$  et  $c$ . On rappelle que  $r$  et  $c$  peuvent être égales à 1 ou être sous la forme  $r = \sum v_j$ ,  $c = \sum v_k / \sum v_i$ . Puisque  $X \in (\mathbb{R}^{>0})^n$ , on aura  $r, c \in \mathcal{C}^1$ . On a aussi que l'ensemble des fonctions  $f(e)$  sont continues sur  $t$  sauf aux points  $d_k \in D$ .

Pour chaque

$$J_j = (d_j, d_{j+1}), \quad (4.1.4)$$

incluant

$$J_0 = (-\infty, d_1) \quad \text{et} \quad J_M = (d_M, \infty), \quad (4.1.5)$$

on prend  $\Omega_j = J_j \times (\mathbb{R}^{>0})^n$  et puisque  $f$  est  $\mathcal{C}^1$  partout sur  $\Omega_j$ , on aura une solution  $X_j$  définie sur  $J_j$  qui sera unique. On peut finalement définir  $\forall t \in \mathbb{R} \setminus D$  la solution

$$X(t) = X_j(t) \text{ si } t \in J_j \quad (4.1.6)$$

qui sera unique sur  $(\mathbb{R} \setminus D) \times (\mathbb{R}^{>0})^n$ . ■

*Corollaire 4.1.2.* Soit  $\Omega = \mathbb{R} \times (\mathbb{R}^{>0})^n$  un ouvert, où  $n = |V(G)|$ . Soit  $f : \Omega \rightarrow \mathbb{R}^n$ , telle que  $f(t, X)$  est  $\mathcal{C}^1$  par morceaux sur  $t$  et  $\mathcal{C}^1$  sur  $X$ . Étant donnée une condition initiale  $(t_0, x_0) \in \Omega$ ,  $X' = f(t, X)$  admet une solution continue  $X^*$  unique sur  $\Omega$ .

DÉMONSTRATION. On pose

$$X^*(t) = \int_{t_0}^t f(u, X(u)) du. \quad (4.1.7)$$

Puisque la fonction  $f(t, X(t))$  est bornée, on sait que

$$\begin{aligned} |X^*(x) - X^*(y)| &= \left| \int_y^x f(u, X(u)) du \right| \leq \left| \int_y^x |f(u, X(u))| du \right| \\ &\leq \left| \int_y^x M du \right| = M|x - y|. \end{aligned} \quad (4.1.8)$$

Soit  $\varepsilon > 0$ . Si on prend  $|x - y| < \varepsilon/M$ , on aura  $|X^*(x) - X^*(y)| < \varepsilon$  et donc la fonction  $X^*$  est continue. On voit également que  $X^*$  est une solution du système, puisque

$$\frac{\partial X^*}{\partial t} = \frac{\partial}{\partial t} \int_{t_0}^t f(u, X(u)) du = f(t, X(t)). \quad (4.1.9)$$

Finalement, si on prend dans la preuve précédente les conditions initiales

$$x_j = \int_{t_0}^{t_j} f(u, X(u)) du = X^*(t_j), \quad (4.1.10)$$

on aura  $X_j = X^*$  lorsqu'on se restreint au domaine  $J_j$ , par le théorème d'unicité des solutions. On aura donc que  $X^*$  est une solution unique sur  $(\mathbb{R} \setminus D) \times (\mathbb{R}^{>0})^n$ , et donc aussi sur  $\Omega$  par continuité de  $X^*$  (puisque  $D$  n'est pas dense dans  $\mathbb{R}$ ). ■

**Remarque 4.1.3.** La condition  $v(t_0) > 0$  pour tout  $v \in V(G)$  est très restrictive. Par exemple, on initie habituellement le modèle *SIR* avec  $R(0) = 0$ , et c'est aussi le cas de plusieurs autres modèles. Cependant, l'utilisation d'un ouvert dans la preuve est inévitable. De plus, si on permet aux nœuds d'être négatifs ( $v(t) > -\epsilon$ ), la fonction  $c$  ne sera plus  $\mathcal{C}^1$  sur  $X$  puisque le terme  $\sum_i v_i$  peut être nul.

Cela dit, il est peut-être possible d'étudier la frontière de  $\Omega$  sur la variable  $X \neq 0$ , c.-à-d.

$$\partial\Omega|_{X \neq 0} = \{X \in (\mathbb{R}^+)^n \setminus \{0\} : \exists i : x_i = 0\} \quad (4.1.11)$$

pour montrer que les solutions qui commencent dans cette région seront aussi uniques. Pour les besoins de ce mémoire, on prendra comme acquis que cela est vrai. Pour la résolution des EDOs au niveau informatique, il ne s'agit pas de trouver la solution générale qui respecte l'équation  $X' = f(t, X)$ , mais plutôt de trouver itérativement une solution particulière  $X$  à partir de l'évaluation de  $f$ . Si  $f$  est définie partout et bornée, on trouvera toujours la solution  $X^*$  (modulo les erreurs numériques) peu importe le point initial

$$(t_0, x_0) \in \mathbb{R} \times (\mathbb{R}^+)^n. \quad (4.1.12)$$

On utilisera donc le fait qu'il existe une solution continue unique  $X^*$  sur  $\Omega = \mathbb{R} \times (\mathbb{R}^+)^n$  pour toute condition initiale  $(t_0, x_0)$  dans  $\Omega$ .

## 4.2. Structure du graphe modifié

**Theorème 4.2.1.** *Soient  $G$  un carquois épidémiologique et  $G' = \text{mod}(G)$ .  $G'$  est aussi un carquois épidémiologique.*

**DÉMONSTRATION.** Soit  $G' = (V', A', f') = \text{mod}(G)$ . Il faut vérifier que toutes les propriétés de la définition 2.2.8 s'appliquent sur  $G'$ . On note d'abord que chaque hyperarc créé dans  $G'$  aura le même type que celui dans  $G$  qui a mené à sa création (voir définition 3.1.3).

- 1)  $(V, A)$  est un carquois. On rappelle qu'un carquois est tel que  $V$  est un ensemble de nœuds et  $A \subseteq \binom{V}{2} \times \Sigma(V)^2$ . Il est trivial que  $V'$  est un ensemble. On doit donc montrer que  $A' \subseteq \binom{V'}{2} \times \Sigma(V)^2$ . Soit  $e' = (u', v', v'_r, v'_c) \in A'$ . On doit montrer que :
- a)  $u' \in V'$       b)  $v' \in V'$       c)  $u' \neq v'$       d)  $v'_r \in \Sigma(V')$       e)  $v'_c \in \Sigma(V')$

On utilise les types d'hyperarcs pour simplifier cette preuve :

- a) Pour les premier et second points, on voit facilement que pour tout  $u'$  ou  $v'$ ,  $u^1$  et  $v^1$  seront dans  $V'$ . Ceci n'est pas nécessairement le cas pour  $u^0$  and  $v^0$ . Cependant  $u^0$  est utilisé pour la création de  $e'$  seulement lorsqu'il existe, et le point (a) est donc respecté.
- b) Pour  $v^0$ , ce terme apparaît dans  $e'$  pour les **rates** dans la couche 0, mais l'utilisation de  $i_{v'}$  prévient son utilisation dans le cas où le nœud n'existe pas. Cette prévention n'est pas nécessaire pour les contacts dans la couche 0 puisque  $v$  sera toujours créé dans la couche 0. Le point (b) est donc respecté.
- c) Si  $u \neq v$ , il est trivial que  $u^i \neq v^j$  pour n'importe quel  $i$  et  $j$ . Le point (c) donc respecté, puisque  $G$  respecte cette propriété.
- d) Pour le point (d), dans le cas des **rates** et des **contacts**, il est clair que  $v_r = u$  et donc  $v_r \in V' \subseteq \Sigma(V')$ . Pour les hyperarcs de type **birth**, l'utilisation de « split » assure que les nœuds de la couche 0 ne sont jamais utilisés s'ils n'existent pas. On voit facilement que  $\text{Null}_n$  et  $v_r^1$  sont dans  $V'$ , ce qui conclue ce point.
- e) Le dernier point est trivial pour les **rates**, **births** et **batches**. Pour les **contacts**, l'utilisation de « split » assure le bon fonctionnement. Pour les hyperarcs  $\mathcal{R}_t$ ,  $v_c^0$  existe toujours puisque tous les nœuds dans  $v_c$  sont dupliqués pour les hyperarcs **contacts**. Le point (e) est donc respecté.
- 2) *Chaque nœud est une fonction du temps.* Cette propriété découle du fait que  $G$  la respecte aussi.
- 3)  $f : A(G) \rightarrow \mathcal{L}^+[t]$ . Cette propriété est respectée, puisque tous les paramètres  $f$  de  $G'$  sont définis à partir de  $G$ , qui respecte cette propriété.
- 4)  $\exists v \in V(G) : v$  est une source. Dans  $G$ , on sait que  $\text{Null}_d$  est une source par la propriété 5 (le nœud  $\text{Null}_d$  ne peut pas être une destination pour un hyperarc). Il n'y a donc pas de  $w \neq \text{Null}_d$  tel que, dans  $G'$ ,  $w^1$  ou  $w^0 = \text{Null}_d$ , ce qui signifie que  $G'$  respecte cette propriété aussi.
- 5)  $\exists v \in V(G) : v$  est un puits. L'argument est semblable au point 4, pour  $\text{Null}_n$ .
- 6)  $v_r \neq \text{Null}_d$  et  $v_c \neq \text{Null}_n$ . Puisqu'on a toujours  $\text{Null}_n^1 = \text{Null}_n^0 = \text{split}(\text{Null}_n) = \text{Null}_n$ , et de même pour  $\text{Null}_d$ , et on sait aussi que  $v \neq \text{Null} \iff v^1, v^0, \text{split}(v) \neq \text{Null}$ , on peut facilement voir que l'algorithme conserve cette propriété à partir de  $G$ .

- 7)  $v_r \neq \text{Null}_n \implies v_r = u$ . Ce point ne concerne que les **rates** et **contacts**. Pour ces deux types d'hyperarcs, la propriété est respectée pour  $G$ . Lors de la création du carquois modifié, on a  $u = v_r \implies u^0 = v_r^0$  et  $u^1 = v_r^1$  et la propriété est donc respectée.
- 8)  $v_c \neq \text{Null}_d \implies v_r \neq \text{Null}_n$ . Cette propriété n'affecte que les **contacts**. On a déjà cette propriété pour les contacts de  $G$ . Puisque  $v \neq \text{Null}_n \iff v^1, v^0 \neq \text{Null}_n$ , on a donc que les contacts dans  $G'$  auront cette propriété.
- 9) *Tous les chemins entre  $S(G)$  et  $I(G)$  contiennent un **contact**.* Par la forme des hyperarcs de type **contacts** dans la définition 3.1.3, on a

$$\begin{aligned} S(G') &= \{\text{split}(v) \in V(G') : v \in S(G)\} \\ I(G') &= \{\text{split}(v) \in V(G') : v \in I(G)\}. \end{aligned} \tag{4.2.1}$$

Supposons que deux nœuds  $u \in S(G')$  et  $v \in I(G')$  soient connectés dans  $G'$  par un chemin  $u_0 u_1 \cdots u_n$  qui ne contient aucun contact. Ce chemin passe à travers des hyperarcs qui ont le même type que ceux du chemin  $b(u_0) b(u_1) \cdots b(u_n)$  dans  $G$ , et donc ce dernier ne contiendra lui non-plus aucun contact. Or, on sait que  $b(u) \in S(G)$  et  $b(v) \in I(S)$ , et donc un tel chemin ne peut pas exister. ■

### 4.3. Dynamique du modèle modifié

Soit  $X(t)$  la solution continue unique du système  $X' = f(t, X)$  et  $Y(t)$  la solution continue unique du système modifié par la définition 3.1.3,  $Y' = g(t, Y)$ , où on fait déjà la somme des compartiments séparés. Pour montrer que la dynamique du modèle est conservée, on veut montrer que les solutions

$$X(t) = [v_i(t)] \quad \text{et} \quad Y(t) = [\text{split}(v_i)(t)] \tag{4.3.1}$$

sont égales  $\forall t \in \mathbb{R}^{\geq \tau}$ , où  $\tau$  est le moment où on souhaite calculer le nombre de reproduction effectif. Pour ce faire, il faut montrer que

$$1) \forall (t, X) \in D, f(t, X) = g(t, X); \quad 2) X(\tau) = Y(\tau).$$

Avec ces deux propriétés, le théorème d'unicité garantit que  $\forall t \in \mathbb{R}, X(t) = Y(t)$ .

**Theorème 4.3.1.** *Soient un hypergraphe de mouvement  $G$  et sa modification  $G' = \text{mod}(G)$ . On aura  $\forall v \in V(G)^*, v(t) = \text{split}(v)(t)$  (voir équation 3.1.2).*

DÉMONSTRATION. Montrons d'abord que, pour le temps initial  $\tau$ ,  $X(\tau) = Y(\tau)$ , c'est-à-dire  $\forall w \in V(G)^*$ ,  $w(\tau) = \text{split}(w)(\tau)$ . Il faut alors s'intéresser à deux cas :

- 1)  $\text{split}(w) = w^1$  : Si le nœud ne se trouve pas dans la couche 0, c'est qu'il n'est pas le point d'arrivée d'un **contact**, et donc on aura  $\delta_{\text{inf}}(w, \tau) = 0 \implies w^1(\tau) = w(\tau)$ .
- 2)  $\text{split}(w) = w^1 + w^0$  : L'équation 3.2.4 nous donne alors  $\text{split}(w)(\tau) = w^1(\tau) + w^0(\tau) = w(\tau) - \delta_{\text{inf}}(w, \tau) + \delta_{\text{inf}}(w, \tau) = w(\tau)$ .

Il faut maintenant vérifier que  $f(t, X) = g(t, X)$ , et donc qu'à chaque point  $X$  où

$$\forall w \in V(G)^*, w(t) = \text{split}(w)(t) \quad (4.3.2)$$

on aura  $w'(t) = \text{split}(w)'(t)$ . Soit un hyperarc  $e$  de  $A^\pm(w)$ . Montrons que cet hyperarc a été divisé dans  $G'$  vers les hyperarcs  $e^1 \in A^\pm(w^1)$  et (parfois)  $e^0 \in A^\pm(w^0) \cup A^\pm(w^1)$  de manière à créer le-s même-s terme-s dans les EDOs, c'est-à-dire

$$\Delta_e(t) = \begin{cases} \Delta_{e^1}(t) + \Delta_{e^0}(t) & \text{si } e^0 \text{ existe} \\ \Delta_{e^1}(t) & \text{si } e^0 \text{ n'existe pas.} \end{cases} \quad (4.3.3)$$

On veut aussi vérifier qu'aucun autre hyperarc qui touche à  $V(G)^*$  n'est créé à partir de  $e$ .

Pour cela, considérons les types d'hyperarcs séparément. Il sera utile d'avoir la définition 3.1.3 en main. Commençons par les hyperarcs entrants  $e \in A^-(w)$ .

- 1)  $e = (\cdot, w, \cdot, \cdot) \in \text{batches}$  : On crée simplement l'hyperarc  $e^1 = (\cdot, w^1, \cdot, \cdot) \in A^-(w^1)$  dans  $G'$  et on a  $\Delta_{e^1}(t) = f'(e^1) = f(e) = \Delta_e(t)$ .
- 2)  $e = (u, w, v_r, \cdot) \in \text{rates}$  : Un ou deux hyperarcs sont créées, soient  $e^1 = (u^1, w^1, v_r^1, \cdot)$  qui sera toujours dans  $A^-(w^1)$  et  $e^0 = (u^0, w^{i^0}, v_r^0, \cdot)$  qui n'existera que si  $u^0 \in G'$  et qui sera soit dans  $A^-(w^1)$  ou  $A^-(w^0)$  (si  $w^0$  existe). Considérons les sous-cas :
  - a)  $e^0$  existe : On aura  $\Delta_{e^1}(t) = f(e)v_r^1(t)$  et  $\Delta_{e^0}(t) = f(e)v_r^0(t)$ . Sous l'hypothèse de l'équation 4.3.2, on aura

$$\begin{aligned} \Delta_{e^1}(t) + \Delta_{e^0}(t) &= f(e)(v_r^1 + v_r^0)(t) = f(e) \text{split}(v_r)(t) \\ &= f(e)v_r(t) = \Delta_e(t). \end{aligned} \quad (4.3.4)$$

- b)  $e^0$  n'existe pas : On sait que  $u^0$  n'existe pas, et donc  $v_r^0$  non plus. On aura donc

$$\Delta_{e^1}(t) = f(e)v_r^1(t) = f(e) \text{split}(v_r)(t) = f(e)v_r(t) = \Delta_e(t). \quad (4.3.5)$$

3)  $e = (u, w, v_r, v_c) \in \text{contacts}$  : L'algorithme crée deux ou trois hyperarcs, soient  $e^1 = (u^1, w^1, v_r^1, \text{split}(v_c)) \in A^-(w^1)$ ,  $e^0 = (u^0, w^0, v_r^0, \text{split}(v_c)) \in A^-(w^0)$  si  $u_0 \in V(G')$  et finalement  $(\cdot, \mathcal{R}_t(e), v_r^1, v_c^0)$  qui pourra être ignoré puisqu'il ne touche pas  $V(G)^*$ . Il est à noter que  $w^0$  sera toujours dans  $G'$  pour les hyperarcs  $\text{contacts}$ . On aura encore deux cas à considérer :

a)  $e^0$  existe : On aura

$$\Delta_{e^1}(t) = f(e)v_r^1(t)\frac{\text{split}(v_c)(t)}{N} \quad \text{et} \quad \Delta_{e^0}(t) = f(e)v_r^0(t)\frac{\text{split}(v_c)(t)}{N}. \quad (4.3.6)$$

Sous l'hypothèse de l'équation 4.3.2, on aura

$$\begin{aligned} \Delta_{e^1}(t) + \Delta_{e^0}(t) &= f(e)(v_r^1 + v_r^0)(t)\frac{\text{split}(v_c)(t)}{N} \\ &= f(e)\text{split}(v_r)(t)\frac{\text{split}(v_c)(t)}{N} = f(e)v_r(t)\frac{v_c(t)}{N} = \Delta_e(t). \end{aligned} \quad (4.3.7)$$

b)  $e^0$  n'existe pas : On aura plutôt

$$\begin{aligned} \Delta_{e^1}(t) &= f(e)v_r^1(t)\frac{\text{split}(v_c)(t)}{N} = f(e)\text{split}(v_r)(t)\frac{\text{split}(v_c)(t)}{N} \\ &= f(e)v_r(t)\frac{v_c(t)}{N} = \Delta_e(t). \end{aligned} \quad (4.3.8)$$

4)  $e = (\cdot, w, v_r, \cdot) \in \text{births}$  :  $A'$  ne contient que l'hyperarc  $e^1 = (\cdot, w^1, \text{split}(v_r), \cdot) \in A^-(w^1)$ . On aura

$$\Delta_{e^1} = f(e)\text{split}(v_r) = f(e)v_r = \Delta_e(t). \quad (4.3.9)$$

Considérons maintenant les hyperarcs sortants  $e \in A^+(w)$ . On veut surtout vérifier que  $e^1$  est bien dans  $A^+(w^1)$  et  $e^0 \in A^+(w^0) \cup A^+(w^1)$ . Les hyperarcs  $\text{batches}$  et  $\text{births}$  sont impossibles dans  $A^+(w)$  car  $w \in V(G)^*$  et ces hyperarcs seront toujours sortants du vide.

1)  $e = (w, v, v_r, \cdot) \in \text{rates}$  : Les mêmes hyperarcs sont créés que dans le cas de  $A^-$ .  $e^1$  sera dans  $A^+(w^1)$  et  $e^0 \in A^+(w^0)$  si  $w^0$  existe. Dans les deux cas, on a le même calcul qu'auparavant, et on a terminé.

2)  $e = (w, v, v_r, v_c) \in \text{contacts}$  : Encore une fois, on a un ou deux hyperarcs qui nous intéressent et un ignoré.  $e^1$  sera toujours dans  $A^+(w^1)$  et  $e^0 \in A^+(w^0)$  si  $w^0$  existe. On a encore les mêmes calculs pour  $\Delta$  dans les deux cas, et on a terminé.

On a donc que la définition 3.1.3 maintient la dynamique des carquois épidémiologiques. ■

## 4.4. Normalisation du nombre de cas index

**Theorème 4.4.1.** *La normalisation des infectés à 1 avant ou après la simulation n'affecte pas le calcul de  $\mathcal{R}_\tau$ .*

DÉMONSTRATION. Supposons qu'au moment  $\tau_0$  tous les compartiments  $v$  dans la chaîne de contact de  $G$  soient divisés dans le nouveau graphe  $G'$  avec les valeurs positives

$$v^1(\tau) = v(\tau) - \widehat{\delta}_{\text{inf}}(b(v), \tau) \quad \text{et} \quad v^0(\tau) = \widehat{\delta}_{\text{inf}}(b(v), \tau). \quad (4.4.1)$$

Soit maintenant une nouvelle situation où ces nœuds sont divisés sans normalisation, c.-à-d.

$$v^{1'}(\tau) = v(\tau) - \delta_{\text{inf}}(b(v), \tau) \quad \text{et} \quad v^{0'}(\tau) = \delta_{\text{inf}}(b(v), \tau) \quad (4.4.2)$$

où  $\delta_{\text{inf}} = k \cdot \widehat{\delta}_{\text{inf}}$  avec  $0 < k \in \mathbb{R}$  et une normalisation sur  $\mathcal{R}_t(e)$  est faite avec  $1/k$ . On souhaite montrer que pour tous les hyperarcs  $e \in \text{contacts}$ , la valeur de  $\mathcal{R}(e)_\tau$  n'a pas changé.

Dans la première situation, on aura

$$\mathcal{R}(e)_\tau = \int_\tau^\infty \beta(t) \frac{\text{split}(u)(t)v^0(t)}{N(t)} dt \quad (4.4.3)$$

et pour la deuxième situation,

$$\mathcal{R}(e)_\tau = \frac{1}{k} \int_\tau^\infty \beta(t) \frac{\text{split}(u)'(t)v^{0'}(t)}{N(t)} dt. \quad (4.4.4)$$

Puisque, pour chaque nœud du graphe, on a  $\text{split}(u)(t) = \text{split}(u)'(t) = u(t)$  et  $\beta(t), N(t)$  sont les mêmes, on veut simplement montrer que

$$\int_\tau^\infty \beta(t) \frac{u(t)v^0(t)}{N(t)} dt = \int_\tau^\infty \beta(t) \frac{u(t)v^{0'}(t)/k}{N(t)} dt \quad (4.4.5)$$

ce qui est vrai si  $v^{0'}(t) = kv^0(t)$  pour tout  $v$  dans la chaîne d'infection.

On a déjà l'égalité pour tous les nœuds lorsque  $t = \tau$  et on veut maintenant montrer

$$\frac{dv^{0'}(t)}{dt} = k \frac{dv^0(t)}{dt} \quad (4.4.6)$$

sous l'hypothèse que  $w^{0'}(t) = kw^0(t)$  pour tout  $w$ . Or, on sait que les seuls hyperarcs qui seront dans la couche 0 sont des **rates** et des **contacts**. Ceux-ci seront toujours proportionnels à un ensemble de nœuds dans la couche 0 (voir équation 3.1.3), qui respectent  $w^{0'}(t) = kw^0(t)$ . On a donc l'égalité voulue sur la dérivée, et donc sur les nœuds de la couche 0. ■

## 4.5. Convergence des nœuds $\mathcal{R}_t$

**Theorème 4.5.1.** *Tous les nœuds  $\mathcal{R}_t$  dans un carquois modifié convergent dans  $\mathbb{R}$  ou divergent vers l'infini.*

DÉMONSTRATION. Soit  $G$  un carquois de mouvement et  $G' = \text{mod}(G)$  sa modification. On souhaite évaluer la convergence des nœuds  $\mathcal{R}_t$  dans  $G'$ . Soit  $g(t) = \mathcal{R}_t(e)(t)$  pour un certain  $e \in \text{contacts}(G)$ . On sait qu'à chaque moment

$$\frac{dg}{dt}(t) = f(e)(t) \cdot \text{split}(v_r(e))(t) \frac{v_c(e)^0(t)}{N(G, t)}. \quad (4.5.1)$$

Supposons que  $v_r^1$ ,  $v_r^0$  et  $v_c^0$  sont initialisés avec des valeurs positives (tel que mentionné à la section 3.3). Puisque  $\forall e \in A(G), u \neq \text{Null} \implies v_r = u$ , on sait que les dérivées de chaque nœud  $v \neq \text{Null}_n$  est sous la forme

$$\frac{dv}{dt} = -F_v(t)v + G_v(t) \geq -F_v(t)v \quad (4.5.2)$$

où  $F_v(t) \geq 0$  est une fonction bornée localement et  $G_v(t) \geq 0$ . Supposons que  $F_v(t) \leq K_v$  sur l'intervalle de temps qui nous intéresse. On aura donc

$$\frac{dv}{dt} \geq -F_v(t)v \geq -K_v v \implies v(t) \geq v(0) \cdot e^{-K_v t} \implies v(t) \geq 0. \quad (4.5.3)$$

Tous les nœuds  $v \neq \text{Null}_n$  du modèle sont donc à valeurs positives, ce qui implique  $g'(t) \geq 0$  et donc  $g(t)$  converge vers  $\mathcal{R}_t(e) \in \mathbb{R}$  ou diverge vers  $\infty$  ■

*Corollaire 4.5.2.* Tous les nœuds de la population sont positifs en tout temps s'ils sont initiés à des valeurs positives.

DÉMONSTRATION. Ce corollaire découle directement de l'équation 4.5.3. ■



# Chapitre 5

## Capacité maximale

La considération d'une capacité maximale pour certains compartiments est un aspect important pour une modélisation représentative d'épidémies. Par exemple, considérons le modèle de la figure 5.1, où le compartiment  $H$  représente les infectés qui sont ensuite hospitalisés. Avec le modèle donné, on aura le système d'équations

$$\begin{aligned} S' &= \delta R - \left( \beta_I \frac{I}{N} + \beta_H \frac{H}{N} \right) S \\ I' &= \left( \beta_I \frac{I}{N} + \beta_H \frac{H}{N} \right) S - (h + \gamma) I \\ H' &= hI - rH \\ R' &= \gamma I + rH - \delta R \end{aligned} \tag{5.0.1}$$

où on supposera  $h > r > \gamma$  et  $\beta_H < \beta_I$ . Autrement dit, On représente une situation où les hôpitaux améliorent réellement la situation de la maladie, autant au niveau de la durée d'infection que pour la transmission.

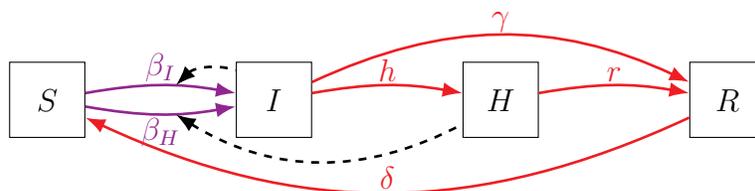


FIGURE 5.1 – Modèle  $SIRS$  avec hospitalisation des infectés.

Avec les valeurs de paramètres

$$\beta_I = 6 \quad \beta_H = 1 \quad \gamma = 1 \quad h = 4 \quad r = 2 \quad \text{et} \quad \delta = 0.1,$$

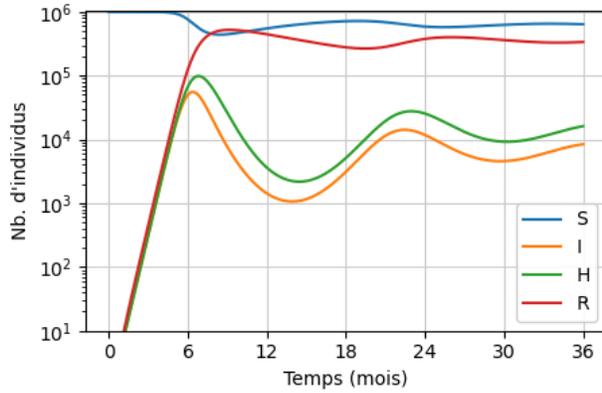


FIGURE 5.2 – Courbes d’infection pour un modèle *SIRS* avec hospitalisation.

jusqu’à ce que les places soient toutes prises. La solution sera donc d’introduire le concept de capacité à l’intérieur des équations du modèle.

## 5.1. Deux types de capacité

**Définition 5.1.1.** La capacité stricte est une manière de contrôler le nombre d’individus en ne laissant pas le nombre total dépasser la capacité donnée. Par exemple, si on considère le nombre de lits d’hôpitaux disponibles, il n’est pas possible de dépasser cette capacité et donc on utilisera une limitation stricte.

**Définition 5.1.2.** La capacité souple permet au nombre d’individus de dépasser la capacité donnée, tout en contrôlant sa croissance vers  $+\infty$ . Par exemple, si on considère le nombre d’individus dans un environnement, il est intuitif que la population ne peut pas être infinie, mais aucun processus strict n’empêche qu’elle devienne très grande pendant un court moment. Dans ce contexte, on parle parfois de « demi-capacité ».

Commençons par expliciter une capacité stricte dans les termes d’équations différentielles. Supposons que l’on souhaite limiter le compartiment  $u$  à une valeur maximale  $u \leq C(u) \in \mathbb{R} \cup \{\infty\}$ . On souhaite simplement que les termes positifs dans l’équation de  $u'(t)$  soient nuls si  $u \geq C(u)$ . Pour ce faire, introduisons une fonction  $c_u : \mathbb{R}^+ \rightarrow [0, 1]$  ( $\mathcal{C}^1$  par morceaux) pour chaque nœud  $u \in V(G)$  qui respecte certaines propriétés (inspirées de [22]) :

- 1)  $\lim_{x \rightarrow 0} c_u(x) = 1$
- 2)  $x \geq C(u) \implies c_u(x) = 0$
- 3)  $\lim_{C(u) \rightarrow \infty} c_u(x) = 1$
- 4)  $\lim_{C(u) \rightarrow 0} c_u(x) = 0$ .

on calcule un  $\mathcal{R}_0^{\text{sim}} \approx 8/5$  et on obtient les courbes simulés de la figure 5.2. On observe ici que  $H > I$  lors des pics d’infection, et que  $H$  atteint un nombre maximal approchant 100 000 individus. Ceci est évidemment un modèle irréaliste qu’il faut adapter pour une modélisation représentative.

Il ne suffit pas de simplement diminuer le paramètre  $h$ , puisqu’on se trouvera parfois dans une situation où on veut accepter le plus d’infectés possible dans les hôpitaux,

Plusieurs choix de  $c_u$  s'offrent à nous pour la capacité stricte, selon les contextes de modélisation. On donne au tableau 5.1 une liste de fonctions pour une capacité stricte ainsi qu'une interprétation possible de leur utilisation.

Fonction	Interprétation
$c_v(x) = \mathbb{1}[x \leq C(v)]$	Les entrées sont contrôlées pour le compartiment seulement lorsqu'on atteint la capacité. Exemple : lits d'hôpitaux sans sélection. Ceci est un cas limite de la 2 <sup>e</sup> fonction, lorsque $n \rightarrow \infty$ .
$c_v(x) = \mathbb{1}[x \leq C(v)] \left(1 - \left(\frac{x}{C(v)}\right)^n\right)$	Les entrées sont plus contrôlées lorsque la taille du compartiment est proche de la capacité. Le « contrôle » est inférieur à $\varepsilon$ lorsque $x \leq \sqrt[n]{\varepsilon}C$ et on a $c_v(C(v)/2) = 1 - (1/2)^n$ . Exemple : lits d'hôpitaux avec sélection, selon la disponibilité.
$c_v(x) = \mathbb{1}[x \leq C(v)] \left(1 - \frac{x}{C(v)}\right)$	Cas spécifique de la fonction précédente. Les entrées sont proportionnelle à la différence entre $C(v)$ et $x$ . On a $c_v(C(v)/2) = 1/2$ .

TABLEAU 5.1 – Exemples de fonctions pour représenter une capacité stricte.

La capacité souple peut être explicitée à l'aide d'une fonction  $c_u : \mathbb{R}^+ \rightarrow [0, 1]$ , qu'on suppose  $\mathcal{C}^1$ , qui satisfait

- |   |  |
|---|--|
| 1) $\lim_{x \rightarrow 0} c_u(x) = 1$      | 4) $\lim_{C(u) \rightarrow \infty} c_u(x) = 1$ |
| 2) $\lim_{x \rightarrow \infty} c_u(x) = 0$ | 5) $\lim_{C(u) \rightarrow 0} c_u(x) = 0$      |
| 3) $\partial_x c_u(x) \leq 0$               | 6) $\partial_{C(u)} c_u(x) \geq 0$ .           |

Le tableau 5.2 illustre quelques choix possibles de  $c_u$  pour une capacité souple.

Fonction	Interprétation
$c_v(x) = \frac{C(v)^n}{C(v)^n + x^n}$	Généralisation d'une formule classique de littérature (voir la prochaine fonction). $c_v(C(v)) = 1/2$ .
$c_v(x) = \frac{C(v)}{C(v) + x}$	Cas particulier de la fonction précédente ( $n = 1$ ). Classiquement utilisée dans la littérature [22]. $c_v(C(v)) = 1/2$ .
$c_v(x) = a^{-x/C(v)}$ avec $a > 1$	Décroissance plus rapide que la fonction précédente à la limite. On a $c_v(C(v)) = 1/a$ .
$c_v(x) = 2^{-x/C(v)}$	Cas spécifique de la fonction précédente. Puisque $a = 2$ , on aura $c_v(C(v)) = 1/2$ .

TABLEAU 5.2 – Exemples de fonctions pour représenter une capacité souple.

## 5.2. Erratum : section 2.3

Une fois qu'on choisit une fonction  $c_u$  pour chaque nœud, on peut ensuite réécrire le système donné à l'équation 3.2.1. Il est à noter que les entrées du modèle ne sont pas directement affectées par la capacité sur les nœuds, mais plutôt par celle sur le modèle en entier (que l'on traitera à la section 5.3). Il faut donc retirer les naissances des termes influencés.

$$\left\{ \begin{array}{l} \frac{dw(t)}{dt} = c_w(w(t)) \sum_{\substack{e^- \in A^-(w) \\ e^- \notin \text{births}}} \Delta_{e^-}(t) + \sum_{\substack{e^- \in A^-(w) \\ e^- \in \text{births}}} \Delta_{e^-}(t) \\ - \sum_{\substack{e^+ \in A^+(w) \\ e^+ \notin \text{births}}} c_{v_{e^+}}(v_{e^+}(t)) \Delta_{e^+}(t) - \sum_{\substack{e^+ \in A^+(w) \\ e^+ \in \text{births}}} \Delta_{e^+}(t) \end{array} \right\} \quad v \in V(G')$$

où  $v_e$  est le nœud d'arrivée de l'hyperarc  $e$ .

Supposons qu'on choisisse les capacités  $C(S) = C(I) = C(R) = \infty$  et  $C(H) = 1000$  et les fonctions  $c_H(H) = \mathbb{1}[H \leq C(H)]$  et  $c_u \equiv 1$  pour les autres nœuds. Le système d'équations 5.0.1 devient alors

$$\begin{aligned} S' &= \delta R - \left( \beta_I \frac{I}{N} + \beta_H \frac{H}{N} \right) S \\ I' &= \left( \beta_I \frac{I}{N} + \beta_H \frac{H}{N} \right) S - (\mathbb{1}[H \leq 1000]h + \gamma)I \\ H' &= \mathbb{1}[H \leq 1000]hI - rH \\ R' &= \gamma I + rH - \delta R \end{aligned} \tag{5.2.1}$$

et on aura  $\forall t, 0 \leq H(t) \leq 1000$ .

Cette information additionnelle de capacité peut être indiquée à l'intérieur du carquois en changeant chaque nœud pour un couple  $(u, C(u))$ . Par exemple, on aurait pour le modèle précédent l'ensemble

$$V(G) = \{(S, \infty), (I, \infty), (H, 1000), (R, \infty), (\text{Null}_n, \infty), (\text{Null}_d, \infty)\}. \tag{5.2.2}$$

Il est à noter que la capacité sur  $\text{Null}_n$  et sur  $\text{Null}_d$  n'a pas de sens (on ne peut pas atteindre une capacité sur les mortalités, et on n'a aucune entrée pour le nœud  $\text{Null}_n$ ), et on dira donc  $C(\text{Null}_n) = C(\text{Null}_d) = \infty$ .

Il est clair que la valeur initiale d'un nœud  $u(0)$  devra être comprise entre 0 et  $C(u)$ . Pour l'information de la fonction à utiliser, il est possible d'inclure dans les fichiers `json`

(voir section 2.5) une chaîne de caractères (*string*), pour chaque compartiment, désignée pour écrire les fonctions  $c_u : \mathbb{R}^+ \rightarrow [0, 1]$  de  $x$ . Dans une interface graphique construite pour un usage général, il serait possible d'inclure un choix de fonctions prédéfinies contenant les fonctions les plus importantes.

### 5.3. Capacité sur population

Pour établir une capacité maximale sur une population entière, on étend la fonction de capacité au domaine  $V(G) \cup \mathcal{G}$  et on choisit une fonction  $c_G$  appropriée (qui respecte les propriétés de capacité stricte ou souple) pour contrôler la population. Les hyperarcs pour lesquels ce contrôle est en question sont ceux qui proviennent de  $\text{Null}_n$ , ce qui laisse place à l'équation 5.3.1. Pour donner la capacité  $C(G)$ , on peut noter le carquois comme un 4-tuple, par exemple  $G = (V, A, f, C(G))$ , ou on peut simplement la noter  $C(G) = C(\text{Null}_n)$ .

Au niveau des équations différentielles, les termes d'entrées sont déjà mis à l'écart et il suffit donc d'ajouter le terme de contrôle :

$$\left\{ \begin{array}{l} \frac{dw(t)}{dt} = c_w(w(t)) \sum_{\substack{e^- \in A^-(w) \\ e^- \notin \text{births}}} \Delta_{e^-}(t) + C_N(N(t)) \sum_{\substack{e^- \in A^-(w) \\ e^- \in \text{births}}} \Delta_{e^-}(t) \\ - \sum_{\substack{e^+ \in A^+(w) \\ e^+ \notin \text{births}}} c_{v_{e^+}}(v_{e^+}(t)) \Delta_{e^+}(t) - C_N(N(t)) \sum_{\substack{e^+ \in A^+(w) \\ e^+ \in \text{births}}} \Delta_{e^+}(t) \end{array} \middle| w \in V(G') \right\} \quad (5.3.1)$$

*Exemple 5.3.1.* Si on ajoute des naissances et mortalités au modèle précédent, avec une capacité de population de  $10^7$  et une fonction  $c_G(x) = \frac{C(G)^{10}}{C(G)^{10} + x^{10}}$ , on peut par exemple obtenir les équations suivantes :

$$\begin{aligned} S' &= A_\alpha \frac{10^{70}}{10^{70} + N^{10}} S + \delta R - \left( \beta_I \frac{I}{N} + \beta_H \frac{H}{N} \right) S - \mu_\alpha S \\ I' &= \left( \beta_I \frac{I}{N} + \beta_H \frac{H}{N} \right) S - (\mathbf{1}[H \leq 1000]h + \gamma) I - (\mu_\alpha + \mu_\beta) I \\ H' &= A_\beta \frac{10^{70}}{10^{70} + N^{10}} H + \mathbf{1}[H \leq 1000]h I - rH - \mu_\gamma H \\ R' &= A_\alpha \frac{10^{70}}{10^{70} + N^{10}} R + \gamma I + rH - \delta R - \mu_\alpha R \end{aligned}$$

Il est à noter que cette notation pour la fonction

$$c_G(N) = \frac{C^n}{C^n + N^n} \quad (5.3.2)$$

mènera éventuellement à des problème numériques pour des grands choix de  $C$  et de  $n$ . Il sera préférable d'implémenter l'écriture

$$c_G(N) = \frac{1}{1 + (N/C)^n}. \quad (5.3.3)$$

## 5.4. Capacité du modèle modifié

Pour respecter la dynamique du modèle original, il devient complexe de définir rigoureusement la capacité de chaque compartiment qui est divisé. Sans rigueur, on peut dire que chaque compartiment divisé verra sa capacité conservée, mais la fonction de capacité sera appliquée sur  $\text{split}(v)$  plutôt que  $v^1$  et  $v^0$  indépendamment puisque l'isolement est virtuel.

Pour les équations différentielles, cela convient parfaitement. C'est plutôt au niveau de la notation, et plus spécifiquement de l'implémentation, qu'on rencontre un problème représentationnel. Une solution serait d'aggréger dans la notation les compartiments qui sont liés, par exemple,

$$V = \left\{ \dots, (\{S\}, C(S)), (\{I^1, I^0\}, C(I)), (\{R\}, C(R)), \dots \right\}. \quad (5.4.1)$$

En adoptant cette notation pour tous les nœuds, on peut alors garder une notation qui a un sens et qui contient l'information nécessaire.

Il est à noter finalement que pour implémenter la capacité maximale, l'ensemble  $A(G)$  devra être modifié pour respecter la notation des nouveaux nœuds. Il faudra également adapter la création des équations différentielles et l'algorithme d'ajout. Cette implémentation n'a pas été réalisée dans le cadre de cette recherche.

# Chapitre 6

---

## Résultats

On compare dans cette section la méthode proposée avec la méthode de la matrice de prochaine génération pour le calcul de  $\mathcal{R}_0$  [7, 24] et avec l'équation 6.1.2 pour le calcul de  $\mathcal{R}_t$  [12]. Ces comparaisons permettent de confirmer l'utilité de notre méthode.

### 6.1. Modèle *SIR*

Un modèle simple qui nous permet de faire une première comparaison est le modèle *SIR* de la figure 6.1. Avec  $f(S, I, S, I) \equiv \beta$  et  $f(I, R, I, \cdot) \equiv \gamma$ , on obtient les équations

$$\begin{aligned} S' &= -\beta S \frac{I}{N} \\ I' &= \beta S \frac{I}{N} + \gamma I \\ R' &= \gamma I. \end{aligned} \tag{6.1.1}$$

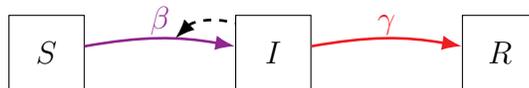


FIGURE 6.1 – Modèle *SIR* simple.

Pour cette version du modèle, on obtient  $\mathcal{R}_0 = \beta/\gamma$ . On peut alors calculer  $\mathcal{R}_t$  à chaque moment en utilisant l'équation donnée en [12], qu'on appellera la « méthode classique », notée

$$\mathcal{R}_t^{\text{ana}} = \mathcal{R}_0 \cdot \frac{S(t)}{N(t)}. \tag{6.1.2}$$

Notre méthode sera notée  $\mathcal{R}_t^{\text{sim}}$ . On rappelle que  $\nu$  représente le nombre de cas incidents.

Pour le modèle *SIR* simple (sans naissances et mortalité), on peut voir que les moments où  $\mathcal{R}_t = 1$  (en orange ---) et  $\nu'(t) = 0$  (en violet ---) sont souvent proches l'un de l'autre, ce qui sera démontré à la section 6.3. Une preuve générale indépendante du modèle n'est pas présentée, mais les résultats de cette section supporteront ce résultat dans certains cas. De plus, on voit déjà dans la figure 6.2 que la méthode classique ne réussit pas très bien à prédire ce maximum d'incidence. On présente dans cette section la courbe  $\nu(t)$  dans tous les graphiques pour comparer la distance entre ces moments pour différents modèles.

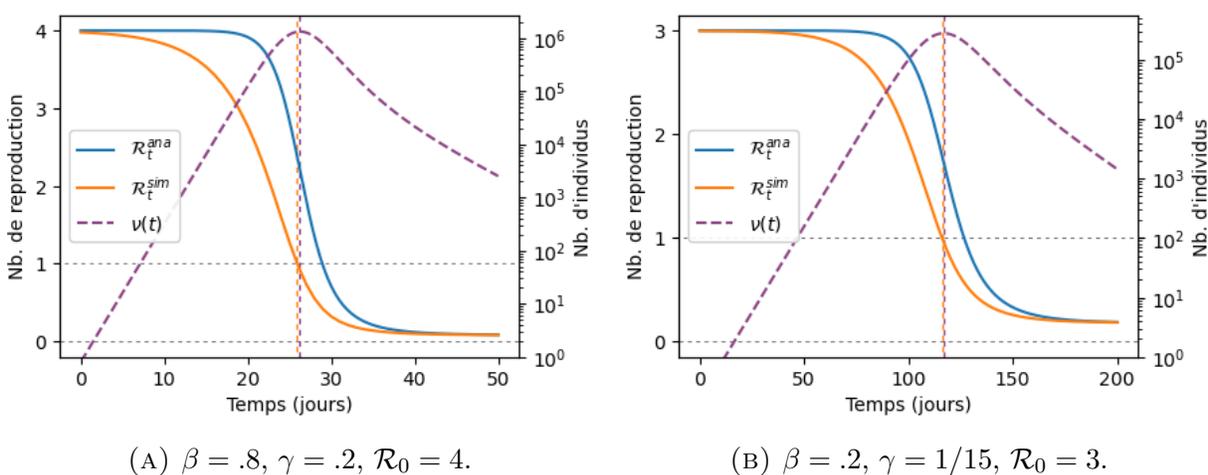


FIGURE 6.2 – Calcul de  $\mathcal{R}_t$  pour deux valeurs de paramètres.

On voit dans la figure 6.2 que notre méthode semble sous-estimer la valeur des nombres de reproductions par rapport à la courbe  $\mathcal{R}_t^{\text{ana}}$ . Contrairement à l'équation 3.2.7 qui est un calcul local, la méthode proposée est une étude globale des contacts dans le modèle. Cela signifie que la diminution de  $S$  au fil du temps affecte directement le calcul de  $\mathcal{R}_0$  et le calcul de  $\mathcal{R}_t$ . De plus, la création de nouveaux infectés dans la couche 1 qui « compétitionnent » avec la couche 0 pour créer de nouvelles infections dans  $S$  influence aussi le calcul.

La dernière chose à noter est que, pour tous les modèles présentés dans cette section, le calcul de  $\mathcal{R}_0^{\text{sim}}$  et  $\mathcal{R}_0^{\text{ana}}$  est très proche. Cela signifie que la méthode de la matrice de prochaine génération produit la même quantité que la méthode explicite présentée, outre la difficulté de sa mise en place et du calcul à effectuer.

## 6.2. Modèles plus complexes

On présente dans cette sous-section quelques modèles dans lesquels certains comportements ne sont pas très bien capturés par le calcul classique  $\mathcal{R}_t = \mathcal{R}_0 \cdot S/N$ . Étant donné un modèle quelconque, cette formule classique ne sera généralement pas appropriée et il faudrait analyser chaque structure de modèle pour modifier la formule adéquatement afin qu'elle s'applique toujours. Cependant, la comparaison entre notre méthode et le calcul simple sert de rappel : la méthode par simulation ne requiert pas de telle modification.

Pour chaque modèle présenté, il sera intéressant de porter attention à la distance entre les extremums des nouvelles infections et les moments où  $\mathcal{R}_t^{\text{sim}} = 1$ , tel que discuté avec le modèle *SIR*.

### 6.2.1. Chaîne de compartiments infectieux

Le premier modèle est celui de la figure A.3, dont les résultats sont montrés à la figure 6.3. Il s'agit d'un modèle *SIR* où la chaîne d'infection est séparée en deux variants. On a cependant un mouvement direct entre  $I_1$  et  $I_2$  et non entre  $I_1$  et  $R$ , ce qui amène une dynamique intéressante. On remarque tout d'abord que la présence de nœuds compteurs pour chaque hyperarc de **contact** nous permet de séparer les courbes  $\mathcal{R}_t$  selon le variant en question, ce qui est utile pour connaître l'évolution de chaque variant dans le temps.

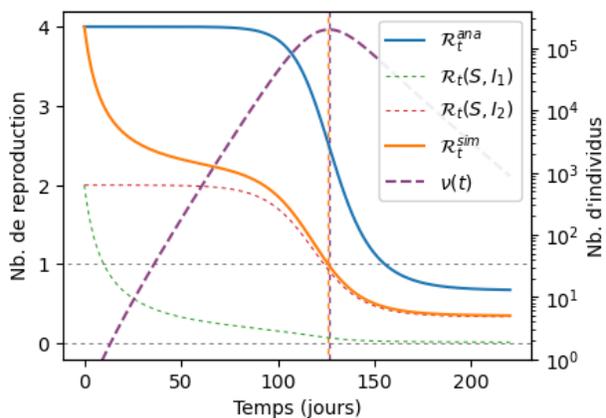
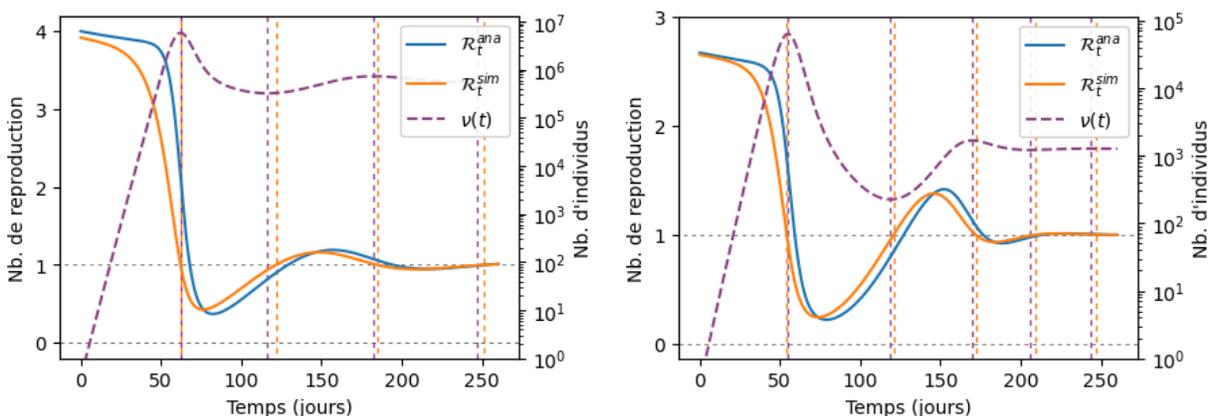


FIGURE 6.3 – Courbes  $\mathcal{R}_t$  pour le modèle de la figure A.3.

### 6.2.2. Phase endémique

On présente à la figure 6.4 deux modèles qui comportent une phase endémique, c'est-à-dire où la maladie atteint un équilibre non nul. En particulier, on étudie à la figure 6.4a le modèle *SIRS* avec vaccination et mortalités présenté à la figure A.5. On voit que tous les extremums d'infection sont capturés par notre méthode de calcul pour  $\mathcal{R}_t$  et de manière un

peu moins précise pour la méthode analytique. On remarque cependant peu de différence entre les deux courbes. Si on enlève la perte d'immunité à l'intérieur de la population, on obtient le modèle *SIR* de la figure A.7 et on observe le même comportement auprès des courbes de reproduction (voir figure 6.4b).



(A) Courbes  $\mathcal{R}_t$  pour le modèle de la figure A.5. (B) Courbes  $\mathcal{R}_t$  pour le modèle de la figure A.7.

FIGURE 6.4 – Courbes  $\mathcal{R}_t$  pour deux modèles avec phase endémique.

### 6.2.3. Séparation de compartiments infectés

Le prochain modèle plus complexe étudié est celui présenté dans [2, 18] et illustré à la figure A.9. Comme discuté dans cet article, la séparation des compartiments infectieux dans un modèle *SIR* mène à une modélisation plus réaliste de la durée d'infection et il sera donc pertinent de s'intéresser à l'impact de cette séparation sur le calcul des nombres de reproduction. On considère dans la figure 6.5 une séparation en trois compartiments *I*.

Tel que discuté à la section 6.1, les nombres de reproduction semblent de manière générale être sous-estimés par le calcul simulé, ce qui est vrai aussi pour le calcul de  $\mathcal{R}_0$ . Or, on voit que cette séparation du compartiment *I* donne une courbe plus élevée au départ, et donc pourrait offrir un calcul plus précis du nombre de reproduction de base. C'est en effet ce qu'on observe avec les données du tableau 6.1.

Nb. compartiments <i>I</i>	1	2	3	4	5
$ \mathcal{R}_0^{\text{ana}} - \mathcal{R}_0^{\text{sim}} $	$2.3 \cdot 10^{-2}$	$3.6 \cdot 10^{-3}$	$7.6 \cdot 10^{-4}$	$2.1 \cdot 10^{-4}$	$8.4 \cdot 10^{-5}$

TABLEAU 6.1 – Convergence de  $\mathcal{R}_0^{\text{sim}}$  selon la séparation du compartiment *I*. Modèle *SIR* simple,  $\beta = 0.8$ ,  $\gamma = 0.2$ .

On remarque également dans la figure 6.5 que le pic d'infection prévu est plus tôt que celui prévu avec le modèle où  $I$  n'est pas séparé. En situation d'épidémie, il est important de ne pas surestimer le temps qui nous est permis pour réagir à une épidémie ou préparer les ressources pour s'en protéger, et on voit ici une indication que le modèle  $SIR$  standard mène à cette erreur.

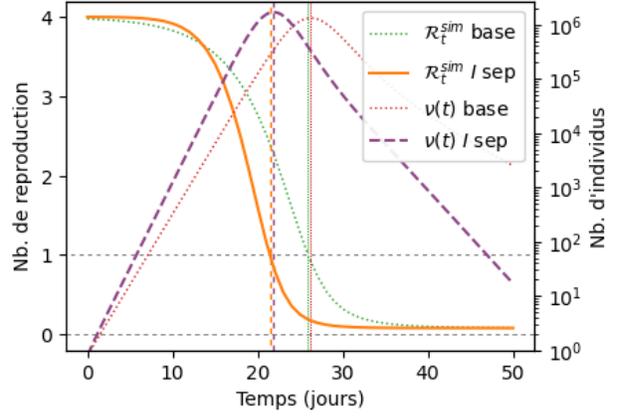


FIGURE 6.5 – Courbes  $\mathcal{R}_t$  pour le modèle de la figure A.9 et le  $SIR$  de base associé.

Il est important de noter que cette modification peut aussi être implémentée de manière automatique dans plusieurs modèles, en utilisant la chaîne d'infection ( $\iota$  dans la définition 3.1.3). En effet, pour tous les compartiments  $w$  de la chaîne d'infection qui ont seulement des sorties de type **rates**, on peut donner une manière plus précise de modéliser le temps passé dans le compartiment en le séparant comme suit :

- 1) Créer les compartiments  $w_1$  à  $w_n$  ;
- 2) Changer tous les hyperarcs entrants de  $w$ , soient ceux sous la forme  $(u, w, v_r, v_c)$ , pour  $(u, w_1, v_r, v_c)$ , en gardant les mêmes paramètres ;
- 3) Soit  $e^+ = (w, v, w, \cdot)$ , un hyperarc sortant de  $w$ . Pour  $1 \leq i \leq (n - 1)$ , créer les hyperarcs  $(e^+)^i = (w_i, w_{i+1}, w_i, \cdot)$ . Créer finalement l'hyperarc  $(e^+)^n = (w_n, v, w_n, \cdot)$  et poser les valeurs de paramètres  $f((e^+)^i) = n f(e^+)$ . Retirer l'hyperarc  $e^+$  ;
- 4) Répéter l'étape 3 pour toutes les autres hyperarcs sortants de  $w$  ;
- 5) Pour tous les hyperarcs où  $w$  fait partie de  $v_c$ , remplacer  $w$  par  $\sum_{1 \leq i \leq n} w_i$  ;
- 6) Retirer le compartiment  $w$ .

Par exemple, pour le modèle  $SIR$  :

$$V(G) = \{S, I, R, \text{Null}_n, \text{Null}_d\} \quad A(G) = \{(S, I, S, I), (I, R, I, \cdot)\}, \quad (6.2.1)$$

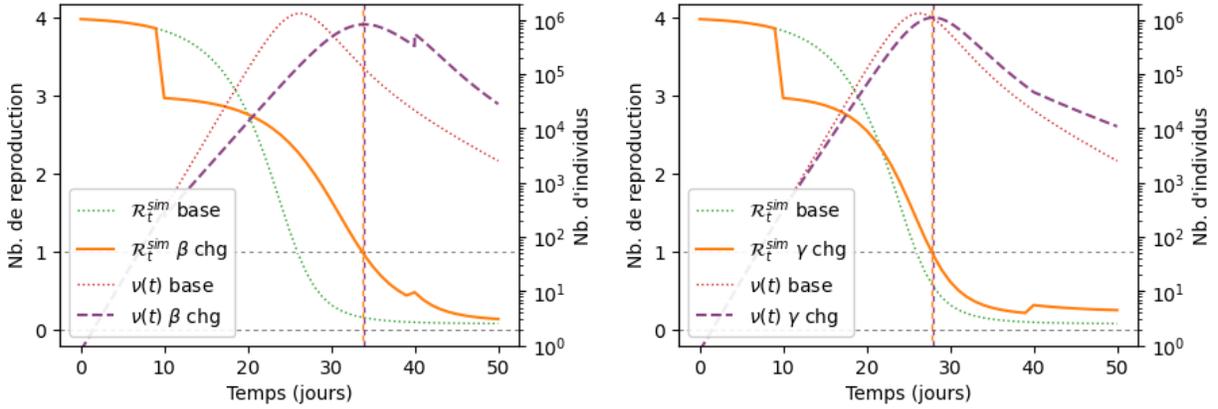
on obtient le nouveau modèle suivant avec  $n$  compartiments  $I$  :

$$V(G) = \{S, I_1, I_2, \dots, I_n, R, \text{Null}_n, \text{Null}_d\} \\ A(G) = \left\{ \left( S, I_1, S, \sum I_i \right), (I_1, I_2, I_1, \cdot), \dots, (I_{n-1}, I_n, I_{n-1}, \cdot), (I_n, R, I_n, \cdot) \right\}. \quad (6.2.2)$$

On peut voir facilement que cette modification du modèle laisse place à un carquois épidémiologique, mais nous n'allons pas vérifier les propriétés formellement. Les étapes décrites sont naturelles et puisqu'on se limite à la chaîne d'infection la dernière propriété (celle qui est la plus stricte) sera respectée.

### 6.2.4. Mesures sanitaires

Le nombre de reproduction  $\mathcal{R}_0$  est affecté directement par les paramètres du modèle tels que le temps d'infection et le taux de contact à l'intérieur de la population. Puisque ces paramètres sont constamment en changement à l'intérieur de populations réelles, on souhaite vérifier que cette évolution ne nuit pas au calcul du nombre de reproduction effectif en appliquant notre méthode. Pour ce faire, nous modifions à différents moments la valeur des paramètres  $\beta$  et  $\gamma$  à l'intérieur d'un modèle *SIR*, ce qui modifie à la fois le nombre analytique  $\mathcal{R}_0^{\text{ana}}$  associé au modèle et le nombre effectif simulé  $\mathcal{R}_t^{\text{sim}}$  en temps réel.



(A)  $\beta$  change de  $4/5 \rightarrow 3/5 \rightarrow 1$ .

(B)  $\gamma$  change de  $1/5 \rightarrow 4/15 \rightarrow 4/25$ .

FIGURE 6.6 – Effet de changements soudains des paramètres sur les courbes de reproduction. La valeur de  $\mathcal{R}_0$  change de 4 à 3 au temps  $t = 10$ , puis de 3 à 5 au temps  $t = 40$ .

Dans la figure 6.6, les paramètres commencent avec les valeurs  $\beta = 4/5$  et  $\gamma = 1/5$ , donc  $\mathcal{R}_0^{\text{ana}} = 4$ . Au temps  $t = 10$ , la valeur de  $\mathcal{R}_0$  est diminuée à 3, soit par la diminution des taux de contacts  $\beta$  (voir figure 6.6a), ou par la diminution du temps de maladie  $1/\gamma$  (voir figure 6.6b). Au temps  $t = 40$ , on modifie à nouveau la valeur du paramètre  $\beta$  ou  $\gamma$  respectivement pour que  $\mathcal{R}_0^{\text{ana}} = 5$ . Cela représente un soulagement des mesures sanitaires après un mois et un retour à une situation qui est moins sécuritaire qu'au départ.

On remarque que le changement se fait de manière beaucoup plus drastique au niveau du paramètre  $\beta$ . Bien que la courbe d'infection atteigne environ la même valeur maximale, elle a été un peu plus étalée dans le temps. Ce changement plus soudain se voit aussi au temps  $t = 40$ , où la courbe d'infection change instantanément dans la figure 6.6a.

### 6.2.5. Modèle à comportement cyclique

On présente finalement un modèle dans lequel les taux de contacts suivent une fonction périodique, montré à la figure A.12. À la figure 6.7, on voit que ce modèle est le premier étudié dans ce mémoire où la courbe générée pour le nombre de reproduction effectif ne prédit pas du tout les extremums de la courbe d'incidence  $\nu$ . En effet, on observe au temps  $t \approx 53$  un minimum d'incidence alors que les deux courbes  $\mathcal{R}_t$  sont à une valeur proche de 1/10. Notons que l'une des prédictions qui est réussie par la méthode proposée, c'est-à-dire au temps  $t \approx 41$ , n'est pas réussie pas la méthode classique qui surestime grandement la valeur du nombre de reproduction effectif.

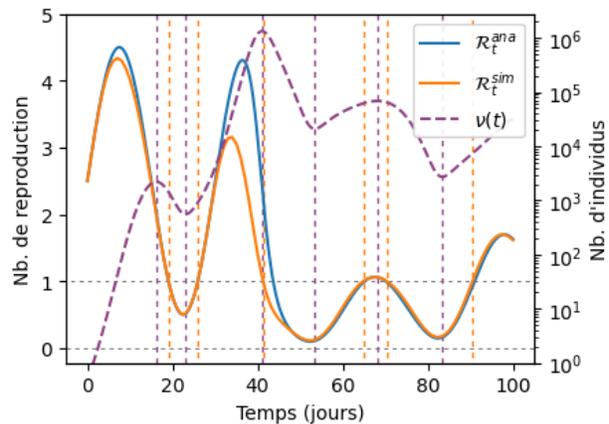


FIGURE 6.7 – Courbes  $\mathcal{R}_t$  pour le modèle de la figure A.12.

## 6.3. Maximum d'incidence totale

Comme le montrent les figures 6.2 et 6.6, on observe avec la majorité des modèles une forte similitude entre le moment où le nombre de nouveaux infectés  $\nu$  atteint un maximum et le moment où  $\mathcal{R}_t$  simulé atteint une valeur de 1. Cette information renforce la capacité de notre modèle à mieux représenter la réalité que la méthode classique, et donc on étudie cette relation sur un modèle *SIR* simple.

Ce résultat ne sera évidemment pas valide sur l'ensemble des modèles épidémiologies construits par la structure de carquois, mais sa validité sur un modèle de base est convaincante pour conclure qu'il vaut la peine d'être étudié.

On prendra le modèle  $SIR$  donné à la section 6.1 :

$$S' = -\beta SI/N \quad I' = \beta SI/N + \gamma I \quad R' = \gamma I \quad (6.3.1)$$

avec les conditions initiales  $S(0) = N - 1$ ,  $I(0) = 1$  et  $R(0) = 0$ . Le calcul classique du nombre de reproduction de base est  $\mathcal{R}_0^{\text{ana}} = \beta/\gamma$  et  $\mathcal{R}_t^{\text{ana}} = \beta/\gamma \cdot S(t)/N(t)$  est celui de l'équation 6.1.2. À partir de l'algorithme de modification de modèle et de calcul de  $\mathcal{R}_t$ , on obtient le modèle modifié montré à la figure A.11, avec

$$\frac{dI^0}{dt} = -\gamma I^0. \quad (6.3.2)$$

Pour calculer  $\mathcal{R}_t$ , on aura  $I^0$  initialisé à  $I^0(t) = \nu(t)$  et donc

$$I^0(u) = \nu(t)e^{-\gamma(u-t)}. \quad (6.3.3)$$

On aura alors

$$\begin{aligned} \mathcal{R}_t^{\text{sim}} &= \frac{1}{\nu(t)} \lim_{\tau \rightarrow \infty} \mathcal{R}_t(S, I)(\tau) = \frac{1}{\nu(t)} \int_t^\infty \beta S^1(u) \frac{I^0(u)}{N} du \\ &= \frac{1}{\nu(t)} \int_t^\infty \beta S(u) \frac{\nu(t)e^{-\gamma(u-t)}}{N} du = \frac{\beta}{N} \int_t^\infty S(u) e^{-\gamma(u-t)} du. \end{aligned} \quad (6.3.4)$$

Pour le modèle présenté, on peut réécrire la dérivée de  $I$  comme ceci :

$$I'(t) = \beta S \frac{I}{N} - \gamma I = \left( \frac{\beta S}{\gamma N} - 1 \right) \gamma I = (\mathcal{R}_t^{\text{ana}} - 1) \gamma I. \quad (6.3.5)$$

Les moments où  $I'(t) = 0$  (extremums du nombre d'infectés au total) sont donc exactement ceux où  $\mathcal{R}_t^{\text{ana}} = 1$ . Cependant, tel que mentionné, on souhaite utiliser les nombres de reproductions calculés par simulation, puisqu'ils sont plus faciles à évaluer pour des modèles plus complexes. On cherche donc un équivalent de cette intersection pour  $\mathcal{R}_t^{\text{sim}}$ .

Dans le modèle  $SIR$  présenté, on aura  $\nu(t) = \beta SI/N$ . Pour trouver les extremums de  $\nu$ , on calcule

$$\begin{aligned} \nu'(t) &= \frac{d}{dt} \frac{\beta SI}{N} = \frac{\beta}{N} \frac{d}{dt} (SI) = \frac{\beta}{N} (S'I + SI') \\ &= \frac{\beta}{N} \left( -\beta S \frac{I}{N} I + S \left( \beta S \frac{I}{N} - \gamma I \right) \right) \\ &= \beta S \frac{I}{N} \left( \frac{\beta}{N} (S(t) - I(t)) - \gamma \right) \end{aligned} \quad (6.3.6)$$

Or, on sait que  $\beta SI/N = \nu$ . De plus, on souhaite obtenir une égalité sous la même forme que l'équation 6.3.5. On aura donc

$$\nu'(t) = \beta S \frac{I}{N} \left( \frac{\beta}{N} (S(t) - I(t)) - \gamma \right) = \nu(t) \gamma \left( \frac{\beta}{\gamma N} (S(t) - I(t)) - 1 \right). \quad (6.3.7)$$

et on définit la nouvelle quantité

$$\mathcal{N}_t = \frac{\beta}{\gamma N} (S(t) - I(t)). \quad (6.3.8)$$

Il est à noter que cette quantité ne remplit pas le même rôle que  $\mathcal{R}_t^{\text{ana}}$  puisqu'elle peut prendre des valeurs négatives lorsque  $I > S$ . Il n'est donc pas clair comment interpréter cette quantité par elle-même. On sait cependant qu'on aura un pic de nouvelles infections exactement lorsque  $\mathcal{N}_t = 1$ . Pour expliquer pourquoi le temps de pic des nouvelles infections et le moment où  $\mathcal{R}_t^{\text{sim}}$  sont aussi proches, il suffit maintenant de voir pourquoi les deux moments où  $\mathcal{R}_t^{\text{sim}} = 1$  et  $\mathcal{N}_t = 1$  sont aussi proches.

**Lemme 6.3.1.** *On a*

$$\mathcal{R}_t^{\text{sim}} = \frac{\beta}{\gamma N} (S(t) - \mathcal{J}_t) \quad (6.3.9)$$

avec

$$\mathcal{J}_t = \int_t^\infty \nu(u) e^{-\gamma(u-t)} du. \quad (6.3.10)$$

DÉMONSTRATION. On intègre par partie le dernier terme obtenu dans l'équation 6.3.4 :

$$\begin{aligned} \mathcal{R}_t^{\text{sim}} &= \frac{1}{N} \int_t^\infty \beta S(u) e^{-\gamma(u-t)} du = \frac{1}{N} \int_{u=t}^\infty \beta S(u) d\left(\frac{e^{-\gamma(u-t)}}{-\gamma}\right) \\ &= \frac{1}{N} \left( \beta S(u) \frac{e^{-\gamma(u-t)}}{-\gamma} \Big|_{u=t}^\infty - \int_{u=t}^\infty \frac{e^{-\gamma(u-t)}}{-\gamma} d(\beta S(u)) \right) \\ &= \frac{1}{N} \left( 0 - \beta S(t) \frac{e^{-\gamma(0)}}{-\gamma} - \int_t^\infty \frac{e^{-\gamma(u-t)}}{\gamma} \beta \nu(u) du \right) \\ &= \frac{1}{N} \left( \frac{\beta S(t)}{\gamma} - \frac{\beta}{\gamma} \int_t^\infty e^{-\gamma(u-t)} \nu(u) du \right) \\ &= \frac{1}{N} \left( \frac{\beta S(t)}{\gamma} - \frac{\beta}{\gamma} \mathcal{J}_t \right) = \frac{\beta}{\gamma N} (S(t) - \mathcal{J}_t). \quad \blacksquare \end{aligned}$$

Le lemme 6.3.1 montre que  $I(t)$  et  $\mathcal{J}_t$  jouent le même rôle respectivement dans le calcul de  $\mathcal{N}_t$  et  $\mathcal{R}_t^{\text{sim}}$ . On a  $\mathcal{R}_t^{\text{sim}} \propto (S(t) - \mathcal{J}_t)$  alors que  $\mathcal{N}_t \propto (S(t) - I(t))$ .

On souhaite maintenant montrer que lorsque  $\nu$  atteint un maximum (i.e.  $\nu'(t) = 0$ ), on aura  $\mathcal{R}_t^{\text{sim}} \approx 1$ . Pour ce faire, on montre que  $\mathcal{J}_t$  et  $I(t)$  sont proches à ce moment.

**Lemme 6.3.2.** *Pour le modèle étudié,*

$$I(t) = e^{-\gamma t} + \int_0^t \nu(u) e^{-\gamma(t-u)} du. \quad (6.3.11)$$

DÉMONSTRATION. On considère la densité de nouveaux rétablis au temps  $t$ , donnée par  $R'(t) = \gamma I(t)$ . On propose un nouveau calcul de cette quantité : on détermine la densité de nouveaux cas créés au temps  $u$  qui se rétablissent au temps  $t > u$ , et on accumule ces rétablissements dans le temps, en ajoutant le premier cas déjà infecté.

Les nouveaux cas qui respectent cette condition sont ceux qui ont une durée d'infection  $d = t - u$ . Puisque la durée d'infection d'un modèle *SIR* suit une exponentielle de distribution  $\gamma$ , la densité sera alors donnée par  $\gamma e^{-\gamma(t-u)}$ . Il suffit alors d'intégrer sur l'ensemble des valeurs de  $u$ . Pour l'individu du départ, on sait qu'il sera au nombre de  $e^{-\gamma t}$  au temps  $t$ , et donc il causera  $\gamma e^{-\gamma t}$  infections au temps  $t$ . On aura alors

$$\begin{aligned} \gamma I(t) &= R'(t) = \int_0^t \nu(u) \gamma e^{-\gamma(t-u)} du + \gamma e^{-\gamma t} \\ &= \gamma \left( \int_0^t \nu(u) e^{-\gamma(t-u)} du + e^{-\gamma t} \right) \\ \implies I(t) &= e^{-\gamma t} + \int_0^t \nu(u) e^{-\gamma(t-u)} du. \quad \blacksquare \end{aligned}$$

Soit le moment  $t^*$  tel que  $\nu'(t^*) = 0 \iff \mathcal{N}_{t^*} = 1$ . Les lemmes 6.3.1 et 6.3.2 nous permettent de calculer

$$\begin{aligned} |I(t^*) - \mathcal{J}_{t^*}| &= \left| e^{-\gamma t^*} + \int_0^{t^*} \nu(u) e^{-\gamma(t^*-u)} du - \int_{t^*}^{\infty} \nu(u) e^{-\gamma(u-t^*)} du \right| \\ &= \left| \int_0^{t^*} \nu(u) e^{-\gamma(t^*-u)} du - \int_{t^*}^{2t^*} \nu(u) e^{-\gamma(u-t^*)} du - \int_{2t^*}^{\infty} \nu(u) e^{-\gamma(u-t^*)} du + e^{-\gamma t^*} \right|. \end{aligned}$$

Par l'inégalité du triangle, on aura

$$\begin{aligned} |I(t^*) - \mathcal{J}_{t^*}| &\leq \left| \int_0^{t^*} \nu(u) e^{-\gamma(t^*-u)} du - \int_{t^*}^{2t^*} \nu(u) e^{-\gamma(u-t^*)} du \right| + e^{-\gamma t^*} + \nu(2t^*) \int_{2t^*}^{\infty} e^{-\gamma(u-t^*)} du \\ &\leq \left| \int_0^{t^*} \nu(u) e^{-\gamma(t^*-u)} du - \int_{t^*}^{2t^*} \nu(u) e^{-\gamma(u-t^*)} du \right| + e^{-\gamma t^*} + \frac{1}{\gamma} e^{-\gamma t^*} \nu(2t^*) \end{aligned}$$

et on peut finalement conclure avec une borne intéressante :

$$|I(t^*) - \mathcal{J}_{t^*}| \leq \underbrace{\left| \int_0^{t^*} \nu(u) e^{-\gamma(t^*-u)} du - \int_{t^*}^{2t^*} \nu(u) e^{-\gamma(u-t^*)} du \right|}_{\text{Mesure de dissymétrie de } \nu \text{ autour de } t^* > 0} + e^{-\gamma t^*} \left( 1 + \frac{\nu(2t^*)}{\gamma} \right) = \Delta \quad (6.3.12)$$

où le deuxième terme est souvent négligeable<sup>1</sup>.

En effectuant certains tests, on réalise que cette mesure de dissymétrie n'est pas si petite. Par exemple, pour les paramètres  $\beta = 2/5$  et  $\gamma = 1/5$ , on obtient une différence entre  $I(t^*)$  et  $\mathcal{J}_{t^*}$  d'environ 4088 (avec le terme négligeable vallant  $2 \cdot 10^{-4}$ ). Cependant, avec la borne déterminée, on peut calculer

$$|\mathcal{R}_{t^*}^{\text{sim}} - \mathcal{N}_{t^*}| = |\mathcal{R}_{t^*}^{\text{sim}} - 1| = \left| \frac{\beta}{\gamma N} (S(t^*) - \mathcal{J}_{t^*}) - \frac{\beta}{\gamma N} (S(t^*) - I(t^*)) \right| = \frac{\beta}{\gamma N} \Delta. \quad (6.3.13)$$

ce qui sera plus acceptable. En recalculant cette erreur avec la borne actuelle, on aura

$$\begin{aligned} |\mathcal{R}_{t^*}^{\text{sim}} - 1| &\leq \frac{\beta}{\gamma N} \left| \int_0^{t^*} \nu(u) e^{-\gamma(t^*-u)} du - \int_{t^*}^{2t^*} \nu(u) e^{-\gamma(u-t^*)} du \right| + \frac{\beta}{\gamma N} e^{-\gamma t^*} \left( 1 + \frac{\nu(2t^*)}{\gamma} \right) \\ &\leq \mathcal{R}_0^{\text{ana}} \left( \left| \int_0^{t^*} \frac{\nu(u)}{N} e^{-\gamma(t^*-u)} du \right| + \left| \int_{t^*}^{2t^*} \frac{\nu(u)}{N} e^{-\gamma(u-t^*)} du \right| \right) + \frac{\mathcal{R}_0^{\text{ana}}}{N} + \mathcal{R}_0^{\text{ana}} \frac{1}{\gamma} \frac{\nu(2t^*)}{N} \end{aligned}$$

Puisque  $S + I \leq N$ , on aura

$$\frac{\nu(u)}{N} = \frac{\beta S I}{N^2} \leq \beta \frac{S(N-S)}{N^2} \leq \frac{\beta}{4} \quad (6.3.14)$$

et on obtient donc

$$\begin{aligned} |\mathcal{R}_{t^*}^{\text{sim}} - 1| &\leq \mathcal{R}_0^{\text{ana}} \cdot 2 \left( \int_0^{t^*} \frac{\beta}{4} e^{-\gamma(t^*-u)} du \right) + \dots \\ &= \mathcal{R}_0^{\text{ana}} \cdot \frac{\beta}{2} \frac{1 - e^{-\gamma t^*}}{\gamma} + \dots \\ &= \frac{1}{2} (\mathcal{R}_0^{\text{ana}})^2 (1 - e^{-\gamma t^*}) + \dots \\ &\leq \frac{1}{2} (\mathcal{R}_0^{\text{ana}})^2 + \dots \end{aligned}$$

On voit évidemment que cette borne est beaucoup trop grande. Il faut donc passer par une autre méthode. Puisque le terme qui se trouve à droite est réduit, travaillons avec seulement

---

1. Pour que ce terme soit plus grand que 1, il faudrait  $\nu(2t^*) \geq \gamma(e^{\gamma t^*} - 1)$ . On rappelle que  $\nu$  atteint un maximum à  $t^*$ , et donc la manière la plus facile que cela arrivera est avec  $\gamma$  très petit

le terme à gauche. On voit que

$$\left| \int_0^{t^*} \frac{\nu(u)}{N} e^{-\gamma(t^*-u)} du - \int_{t^*}^{2t^*} \frac{\nu(u)}{N} e^{-\gamma(u-t^*)} du \right| = \left| \int_0^{t^*} h(x) dx - \int_{t^*}^{2t^*} h(x) dx \right| \quad (6.3.15)$$

avec

$$h(x) = \frac{\nu(x)}{N} \begin{cases} e^{-\gamma(t^*-x)} & \text{si } x \leq t^* \\ e^{-\gamma(x-t^*)} & \text{si } x \geq t^*. \end{cases} \quad (6.3.16)$$

Puisqu'on s'attend à ce que la différence d'intégrale soit petite, il vaut la peine de faire le changement de variable  $u = 2t^* - x \implies x = 2t^* - u$  dans la deuxième intégrale :

$$\begin{aligned} \left| \int_0^{t^*} h(x) dx - \int_{t^*}^{2t^*} h(x) dx \right| &= \left| \int_0^{t^*} h(x) dx - \int_{u=t^*}^0 h(2t^* - u) d(2t^* - u) \right| \\ &= \left| \int_0^{t^*} h(x) dx + \int_{t^*}^0 h(2t^* - u) du \right| = \left| \int_0^{t^*} h(x) dx - \int_0^{t^*} h(2t^* - u) du \right|. \end{aligned} \quad (6.3.17)$$

On peut réécrire  $h(x)$  avec sa valeur et l'équation 6.3.17 devient

$$\begin{aligned} \left| \int_0^{t^*} \frac{\nu(u)}{N} e^{-\gamma(t^*-u)} du - \int_0^{t^*} \frac{\nu(2t^* - u)}{N} e^{-\gamma(t^*-u)} du \right| &= \left| \int_0^{t^*} \frac{\nu(u) - \nu(2t^* - u)}{N} e^{-\gamma(t^*-u)} du \right| \\ &\leq \int_0^{t^*} \left| \frac{\nu(u) - \nu(2t^* - u)}{N} e^{-\gamma(t^*-u)} \right| du \leq \int_0^{t^*} \frac{|\nu(u) - \nu(2t^* - u)|}{N} du \end{aligned}$$

ce qui nous donne

$$\left| \mathcal{R}_{t^*}^{\text{sim}} - 1 \right| \leq \mathcal{R}_0^{\text{ana}} \left( \int_0^{t^*} \frac{|\nu(u) - \nu(2t^* - u)|}{N} du + \frac{1}{N} + \frac{1}{\gamma} \frac{\nu(2t^*)}{N} \right). \quad (6.3.18)$$

Évidemment, dans ce calcul, le terme  $\mathcal{R}_0^{\text{ana}}/N$  n'aura presque pas d'importance, et le terme  $\nu(2t^*)/\gamma$  se comportera plutôt bien si  $\gamma$  n'est pas trop petit. En particulier, pour que

$$\frac{\mathcal{R}_0^{\text{ana}} \nu(2t^*)}{\gamma N} \quad (6.3.19)$$

soit grand, il faudrait que  $\beta/\gamma^2$  soit très petit ou qu'une majorité de la population soit infectée ponctuellement au moment  $2t^*$  (il faut se rappeler que le maximum pour  $\nu$  est dépassé à ce moment).

Pour conclure l'analyse de cette concordance, on montre à la figure 6.8 le calcul réel de  $|\mathcal{R}_{t^*}^{\text{sim}} - 1|$  lorsque  $N = 10^6$ . On note que la différence réelle semble être seulement une fonction de  $\mathcal{R}_0^{\text{ana}}$  lorsque  $N$  est fixe. On note finalement que lorsque  $\mathcal{R}_0^{\text{ana}} > 5$  (par exemple), on aura une différence entre  $\mathcal{R}_{t^*}^{\text{sim}}$  et 1 plus grande que  $10^{-1}$ , ce qui est plutôt élevé. De

manière générale, on peut calculer

$$\frac{d}{dt} \mathcal{R}_t^{\text{sim}} = \frac{d}{dt} \left( e^{\gamma t} \cdot \frac{1}{N} \int_t^\infty \beta S(u) e^{-\gamma u} du \right) = \gamma \mathcal{R}_t^{\text{sim}} - \beta \frac{S(t)}{N} \quad (6.3.20)$$

et conclure qu'une différence  $\delta$  entre  $\mathcal{R}_{t^*}^{\text{sim}}$  et 1 sera acceptable (en supposons une dérivée plutôt constante) lorsque

$$\delta < \left| \gamma \mathcal{R}_{t^*}^{\text{sim}} - \beta \frac{S(t^*)}{N} \right|. \quad (6.3.21)$$

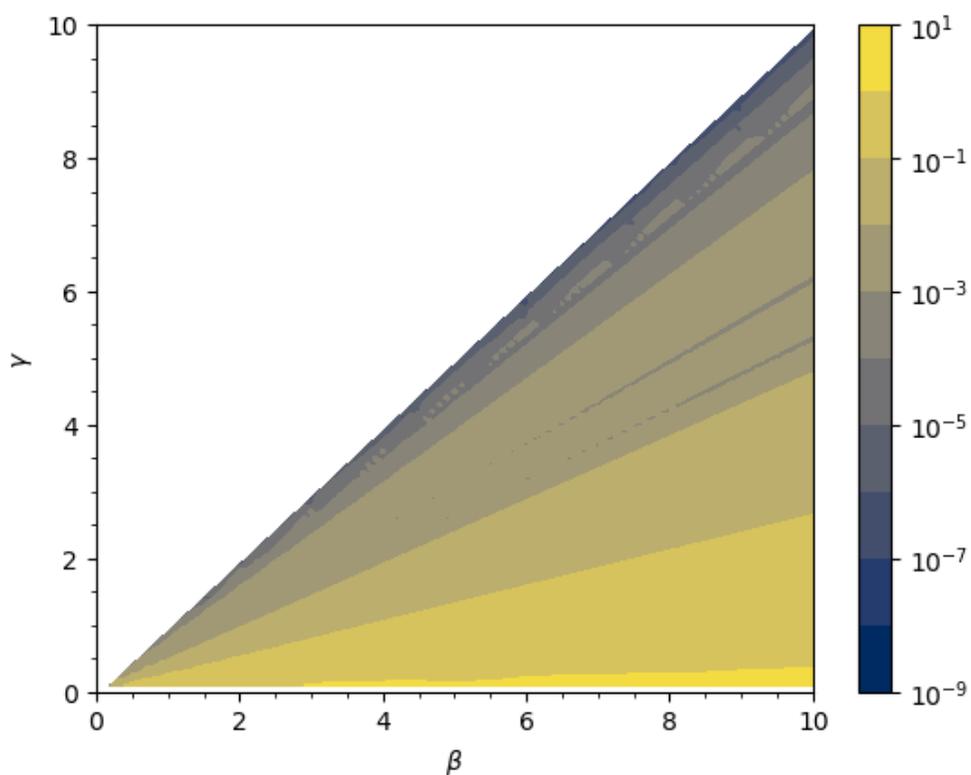


FIGURE 6.8 – Valeur de  $|\mathcal{R}_{t^*}^{\text{sim}} - 1|$  pour un ensemble de valeurs de  $\beta$  et  $\gamma$  entre 0 et 10 telles que  $\mathcal{R}_0 > 1$ , lorsque  $N = 10^6$ .



## Conclusion

---

La structure d'hypergraphes présentée dans ce mémoire est construite pour permettre une formalisation des modèles épidémiologiques, en utilisant un hypergraphe défini soigneusement pour créer les termes voulus dans un système d'EDO. Cette structure amène aussi plusieurs autres avantages pour l'analyse structurelle des modèles. Entre autres, on a montré comment cette nouvelle structure laisse place à une caractérisation des différents types de mouvements entre les compartiments, une définition formelle des compartiments infectés et susceptibles, ainsi qu'une description du mouvement suivi par les infectés. La définition explicite de ces concepts permet à son tour une modification du modèle étudié afin d'isoler certains individus au niveau des équations différentielles. Cet isolement donne finalement une explicitation du calcul des nombres de reproduction qui offre des résultats plus convaincants que la méthode standard  $\mathcal{R}_t = \mathcal{R}_0 S/N$ .

Quelques notions présentes dans la littérature ont également été adaptées à la structure présentée pour permettre une modélisation plus complète, tout en maintenant le calcul des nombres de reproduction. Notamment, on a discuté de

- 1) la fusion de différents modèles épidémiologiques en un seul (à travers le produit cartésien de deux hypergraphes) ;
- 2) la capacité maximale pour contrôler la taille de certaines quantités ;
- 3) la propagation arrière pour l'algorithme de calcul des nombres de reproduction, pour permettre l'apprentissage automatique des paramètres.

Ces notions permettent de créer des modèles qui possèdent des comportements complexes et dynamiques qui n'ont pas été étudiés en détail dans ce mémoire. Par conséquent, des recherches futures pourraient se pencher sur ces ajouts en détail et étudier leur effet sur le calcul des nombres de reproduction.

Tel que discuté à la section 2.5, la création d'équations différentielles à partir des hyperarcs d'un carquois permet un stockage simple de modèles, mais limite aussi la complexité des termes possibles. En écrivant un algorithme de lecture des fonctions plus détaillé, il serait possible de considérer des fonctions plus complexes et ainsi de résoudre cette lacune. Cependant, il faut s'assurer que cette nouvelle écriture permette tout de même une classification utile des hyperarcs de l'hypergraphe, pour qu'une analyse de la structure du modèle soit possible automatiquement. Même si une telle classification n'est pas possible avec une forme plus générale, la lecture de fonctions plus complexes pourrait servir à la capacité maximale, pour stocker des fonctions de capacité qui varient selon plusieurs nœuds.

Le dernier point qui serait intéressant à explorer davantage est la concordance entre les moments où l'incidence des infectés atteint un extremum et ceux où  $\mathcal{R}_t = 1$ . On a démontré la proximité de ces moments pour le modèle *SIR* simple (où  $\mathcal{R}_0$  n'est pas trop grand), mais ce résultat semble être valide pour plusieurs autres modèles plus complexes, comme on peut voir à la section 6.2. Il serait pertinent d'étudier les conditions qui assurent que le calcul proposé aura cette propriété.

Malgré ces quelques omissions, on a tout de même réussi à construire une structure de modèle qui allège le travail et le temps requis pour le calcul des nombres de reproduction, ainsi que pour la construction et la modification de modèles. En offrant une implémentation informatique en Python de la méthode décrite, on espère que le processus de modélisation épidémiologique sera accéléré et que les résultats futurs liés aux modèles à compartiments seront plus précis, tout en demandant un effort moindre. Finalement, le partage de modèles entre les modélisateurs se verra facilité par le stockage de ceux-ci à l'intérieur de fichiers `json` et on souhaite donc que la collaboration dans le domaine de l'épidémiologie mathématique soit mise de l'avant.

## Apprentissage automatique des paramètres

Dans le cadre de ce projet, l'implémentation de l'algorithme d'apprentissage automatique n'est pas faite dans le code, puisque la partie du mémoire portant sur la formalisation des graphes s'est avérée être suffisante pour l'étendue de la maîtrise. Cependant, je juge important de décrire ici une méthode que j'ai en tête afin de conclure cette recherche, basée sur la librairie `torchdiffeq` [4].

On supposera pour cette section que tous les paramètres  $f(e)$  des hyperarcs du carquois  $G$  sont des fonctions constantes, c.-à-d.  $f(e)(t) \equiv f_e$ , et notons les paramètres à apprendre

$$f = \begin{bmatrix} f_{e_1} & f_{e_2} & \cdots & f_{e_n} \end{bmatrix}. \quad (6.3.22)$$

Supposons qu'on ait observé une courbe de nombre de reproduction  $\mathcal{R}_t^*$  (disons, celle donnée par INSPQ, dans la cas de la COVID-19<sup>2</sup>) et qu'on veuille calibrer les paramètres du modèle  $G$  pour que les nombres de reproduction simulés respectent cette courbe (dans la mesure du possible). Puisque la librairie torchdiffeq nous permet de faire la propagation arrière pour les solutions d'un système d'EDOs en fonction des paramètres initiaux, on pourra donc connaître le gradient de chaque  $\mathcal{R}_t$  calculé pour le modèle  $G$  en fonction des paramètres des équations différentielles. On sépare ici l'étude de la méthode analytique  $\mathcal{R}_t = \mathcal{R}_0 \cdot S/N$  et la méthode par simulation proposée.

### Méthode classique basée sur $\mathcal{R}_0$

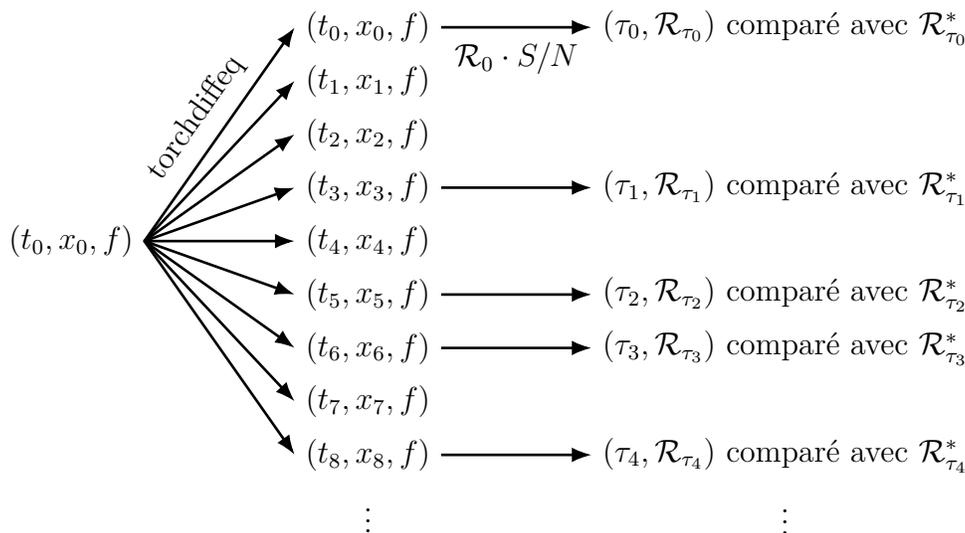


FIGURE 6.9 – Modèle de prédiction pour les nombres de reproduction à partir d'une condition initiale et des paramètres  $f$ .

Pour la première étape de simulation (la partie gauche de la figure 6.9), on trouve les solutions du système pour chaque point discret  $t_0$  à  $t_n$  et, pour chaque point  $x_0$  à  $x_n$  de la solution, la fonction odeint de torchdiffeq nous permet de calculer  $\nabla_f x_i$ .

2. <https://www.inspq.qc.ca/covid-19/donnees>

Pour le calcul des nombres de reproduction effectifs (la partie droite de la figure 6.9), on calcule simplement pour un temps  $\tau_j$  le nombre de susceptibles dans le modèle et la population totale, puis on calcule  $\mathcal{R}_{\tau_j} = \mathcal{R}_0 \cdot S(\tau_j)/N(\tau_j)$ . On aura alors

$$\nabla_f \mathcal{R}_\tau = \frac{S(\tau)}{N(\tau)} \cdot \nabla_f \mathcal{R}_0 + \frac{\mathcal{R}_0}{N(\tau)^2} (N(\tau) \nabla_f S(\tau) - S(\tau) \nabla_f N(\tau)). \quad (6.3.23)$$

En récoltant ces gradients pour chaque  $\mathcal{R}_\tau$ , on peut utiliser une fonction d'erreur (*loss function*) sur l'ensemble des valeurs  $\mathcal{R}_\tau$  en comparant avec  $\mathcal{R}_\tau^*$  et minimiser cette fonction avec un algorithme de descente de gradient (ou tout autre algorithme d'optimisation) en fonction des paramètres  $f$ . Si on le souhaite, `torchdiffeq` nous permet aussi de calculer les dérivées en fonction de la condition initiale  $x_0$ , et on peut ainsi apprendre la condition initiale optimale.

Cependant, ce processus d'optimisation dépend du nombre de reproduction de base  $\mathcal{R}_0$ , qu'il faut calculer avec la matrice de prochaine génération. Comme il a été mentionné quelques fois, l'utilisation de cette méthode implique une analyse mathématique propre à chaque modèle, ce qu'on tente d'éviter.

## Méthode par simulation explicite

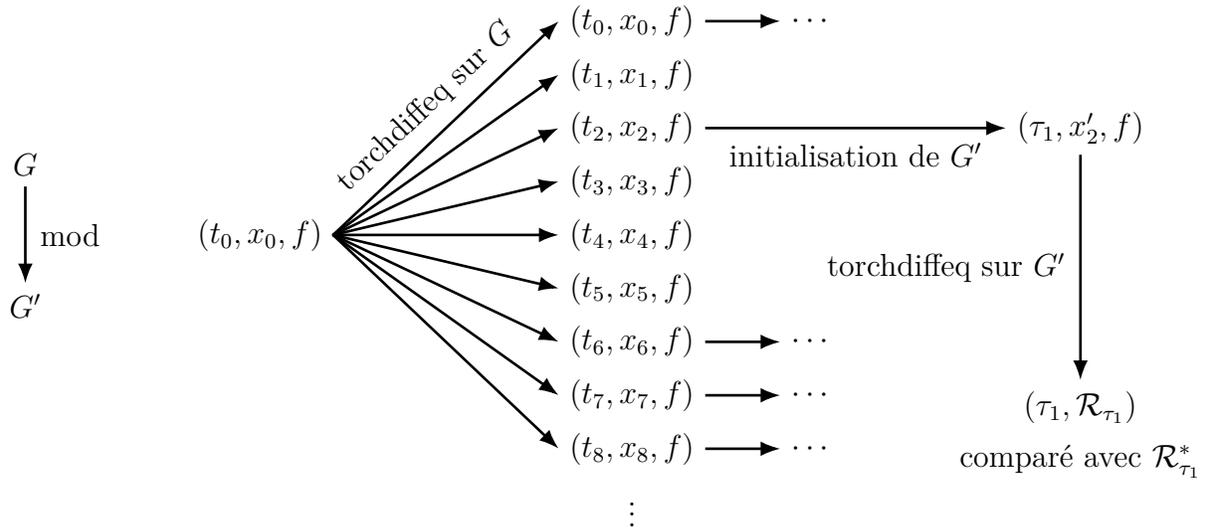


FIGURE 6.10 – Modèle de prédiction pour les nombres de reproduction à partir d'une condition initiale et des paramètres  $f$ , pour la méthode proposée.

Pour la méthode proposée, on commence tout de même par une première étape de simulation (la partie gauche de la figure 6.9) où `torchdiffeq` peut calculer  $\nabla_f x_i$ . Pour le calcul

des nombres de reproduction effectifs, on aura à partir d'une sélection de temps  $\tau_0$  à  $\tau_m$  un ensemble de nouvelles solutions du système basées sur les points  $x_{i_0}$  à  $x_{i_m}$ . Cela équivaut à remplacer le calcul  $\mathcal{R}_0 \cdot S/N$  par une autre étape effectuée par torchdiffeq, voir figure 6.10.

Cependant, ce ne sont que les nœuds  $\mathcal{R}_t(\cdot)$  qui nous intéressent, en particulier la valeur de ces nœuds à la fin de la simulation. On peut calculer le nombre de reproduction total effectif avec la somme de ces valeurs, et on obtient

$$\nabla_f \mathcal{R}_{\tau_j} = \sum_{e \in \text{contacts}} \nabla_f \mathcal{R}_t(e)(T) \quad (6.3.24)$$

où  $T$  est le temps final de simulation. Si on souhaite faire un apprentissage sur  $x_0$  également, on peut calculer

$$\left[ \nabla_f \mathcal{R}_{\tau_j} \quad \nabla_{x_0} \mathcal{R}_{\tau_j} \right] = \sum_{e \in \text{contacts}} \left[ \nabla_f \mathcal{R}_t(e)(T) \quad \nabla_{x_0} x_{i_j} \cdot \nabla_{x_{i_j}} \mathcal{R}_t(e)(T) \right], \quad (6.3.25)$$

où  $x_{i_j}$  est la condition initiale pour le calcul de  $\mathcal{R}_{\tau_j}$ . Comme pour la méthode analytique, on choisit une fonction d'erreur entre les prédictions du modèle et les observations réelles, et on peut ainsi effectuer un processus d'optimisation sur les paramètres  $f$ .

En utilisant une structure de modèle assez complexe (pour créer une dynamique suffisamment complexe au niveau de la courbe  $\mathcal{R}_t$ ), on peut espérer la convergence du vecteur  $f$  vers les paramètres idéaux  $f^*$ , qui minimisera la distance entre la courbe calculée et la courbe observée. En particulier, si on arrive à obtenir la courbe de reproduction de manière assez précise, on peut soupçonner que le modèle utilisé est utile pour prédire la croissance de l'épidémie modélisée.

Cependant, il faut éviter de conclure ainsi trop hâtivement. En effet, avec une structure trop complexe, il sera possible de faire du sur-ajustement de paramètres (overfitting) et obtenir une courbe de  $\mathcal{R}_t$  qui respecte très bien les observations, mais qui est un mauvais prédicteur des nombres de reproduction futurs. Trouver les bons paramètres pour à la fois respecter la courbe de référence et donner de bonnes prédictions pourrait s'avérer impossible par le processus décrit.



# Références bibliographiques

---

- [1] Nithya C ACHAI AH, Sindhu B SUBBARAJASETTY et Rajesh M SHETTY :  $R_0$  and  $R_e$  of COVID-19 : Can we predict when the pandemic outbreak will be contained? Indian Journal of Critical Care Medicine, 24(11):1125–1127, 2020.
- [2] Luca BOLZONI, Rossella DELLA MARCA et Maria GROPPI : On the optimal control of SIR model with Erlang-distributed infectious period : Isolation strategies. Journal of Mathematical Biology, 83(4), 2021.
- [3] Marc BRISSON : Modélisation de l'évolution de la COVID-19 au Québec [Présentation]. 24e journées annuelles de santé publique, 2021. URL : [https://marc-brisson.net/covid19-response/Presentation\\_JASP.pdf](https://marc-brisson.net/covid19-response/Presentation_JASP.pdf).
- [4] Ricky T. Q. CHEN : torchdiffeq, 2018. URL : <https://github.com/rtqichen/torchdiffeq>.
- [5] Ricky T. Q. CHEN, Yulia RUBANOVA, Jesse BETTENCOURT et David DUVENAUD : Neural ordinary differential equations, 2019.
- [6] O. DIEKMANN, J. A. P. HEESTERBEEK et J. A. J. METZ : On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. Journal of Mathematical Biology, 28(4), 1990.
- [7] O. DIEKMANN, J. A. P. HEESTERBEEK et M. G. ROBERTS : The construction of next-generation matrices for compartmental epidemic models. Journal of The Royal Society Interface, 7(47):873–885, 2009.
- [8] Ramsès DJIDJOU-DEMASSE, Christian SELINGER et Mircea T. SOFONEA : Épidémiologie mathématique et modélisation de la Pandémie de Covid-19 : Enjeux et Diversité. Revue Francophone des Laboratoires, 2020(526):63–69, 2020.
- [9] Geoffrey P. GARNETT : The geographical and temporal evolution of sexually transmitted disease epidemics. Sexually Transmitted Infections, 78(Supplement 1):i14–i19, 2002.
- [10] Geoffrey P. GARNETT : Role of herd immunity in determining the effect of vaccines against sexually transmitted disease. The Journal of Infectious Diseases, 191(Supplement 1), 2005.
- [11] GEEKSFORGEEKS : Print all paths from a given source to a destination. GeeksforGeeks. URL : <https://www.geeksforgeeks.org/find-paths-given-source-destination/> (version : 23 janv. 2023).
- [12] Katelyn M. GOSTIC, Lauren MCGOUGH, Edward B. BASKERVILLE et al. : Practical considerations for measuring the effective reproductive number,  $R_t$ . PLOS Computational Biology, 16(12), 2020.

- [13] J. M. HEFFERNAN, R. J. SMITH et L. M. WAHL : Perspectives on the basic reproductive ratio. Journal of The Royal Society Interface, 2(4):281–293, 2005.
- [14] M. W. HIRSCH, S. SMALE et R. L. DEVANEY : Differential Equations, Dynamical Systems, and an Introduction to Chaos. Elsevier, 2004.
- [15] Thomas V. INGLESBY : Public health measures and the reproduction number of SARS-COV-2. JAMA, 323(21):2186, 2020.
- [16] James Holland JONES : Notes on  $R_0$ . Departments of Anthropological Sciences, Standord University. URL : <https://web.stanford.edu/~jhj1/teachingdocs/Jones-on-R0.pdf>.
- [17] Matt J. KEELING et Pejman ROHANI : Modeling infectious diseases in humans and animals. Princeton University Press, 2008.
- [18] Olga KRYLOVA et David J. EARN : Effects of the infectious period distribution on predicted transitions in childhood disease dynamics. Journal of The Royal Society Interface, 10(84):20130098, 2013.
- [19] Ying LIU et Joacim ROCKLÖV : The effective reproductive number of the Omicron variant of SARS-COV-2 is several times relative to Delta. Journal of Travel Medicine, 29(3), 2022.
- [20] Mathieu MAHEU-GIROUX, Arnaud GODIN et Yiqing XIA : Taux de reproduction ( $R_t$ ), des cas de SRAS-CoV-2 au Québec. INSPQ. URL : <https://www.inspq.qc.ca/covid-19/donnees/rt> (version : 29 mars 2023).
- [21] Chris RACKAUCKAS : Comparing Algorithmic Complexity, ODE Solvers (Big O). Computational Science Stack Exchange. URL : <https://scicomp.stackexchange.com/q/28619> (version : 19 juill. 2020).
- [22] Chunhua SHAN et Huaiping ZHU : Bifurcations and complex dynamics of an SIR model with the impact of the number of hospital beds. Journal of Differential Equations, 257(5):1662–1688, 2014.
- [23] P. van den DRIESSCHE et James WATMOUGH : Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Mathematical Biosciences, 180(1-2):29–48, 2002.
- [24] Pauline van den DRIESSCHE : Reproduction numbers of infectious disease models. Infectious Disease Modelling, 2(3):288–303, 2017.
- [25] Lee WORDEN et Travis C. PORCO : Products of compartmental models in epidemiology. Computational and Mathematical Methods in Medicine, 2017:1–16, 2017.

# Annexe A

---

## Modèles additionnels et modifications

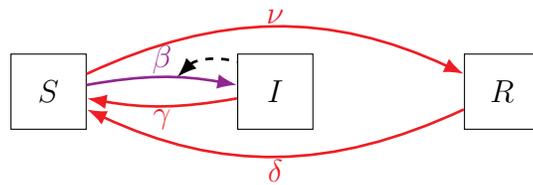


FIGURE A.1 – Modèle  $SIS$  avec vaccination.

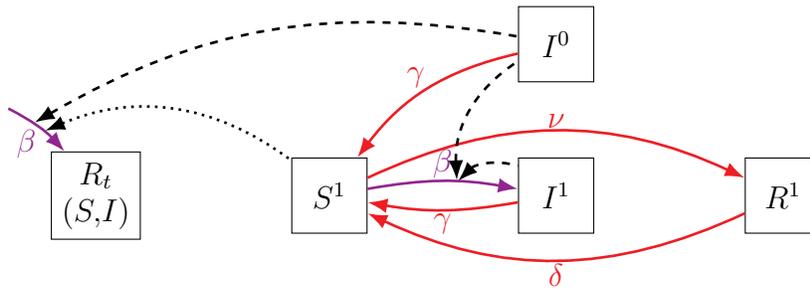


FIGURE A.2 – Modification du modèle de la figure A.1.

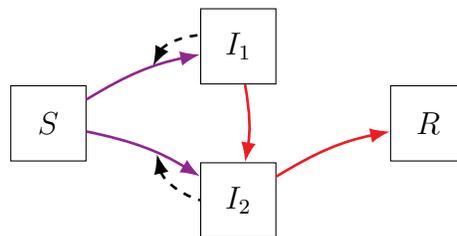


FIGURE A.3 – Modèle  $SIR$  avec deux étapes d'infection.

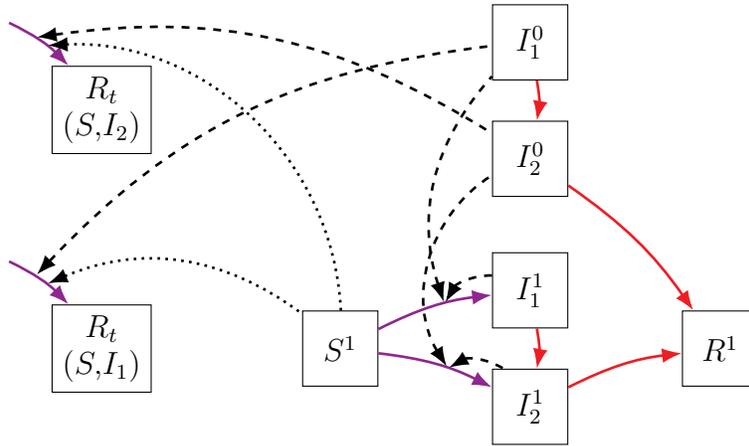


FIGURE A.4 – Modification du modèle de la figure A.3.

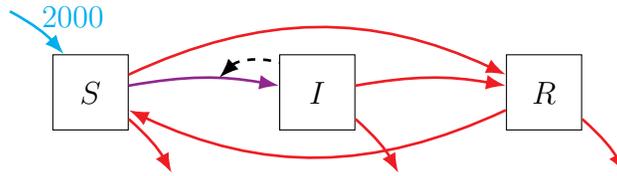


FIGURE A.5 – Modèle *SIRS* avec démographie et vaccination.

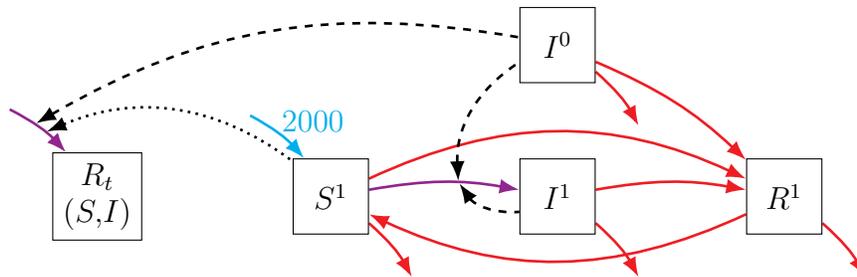


FIGURE A.6 – Modification du modèle de la figure A.5.

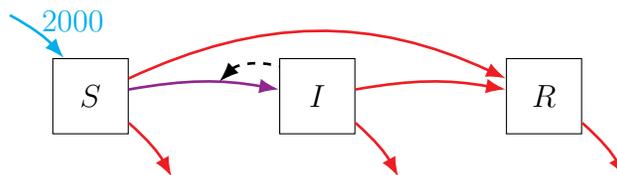


FIGURE A.7 – Modèle *SIR* avec mortalité et vaccination.

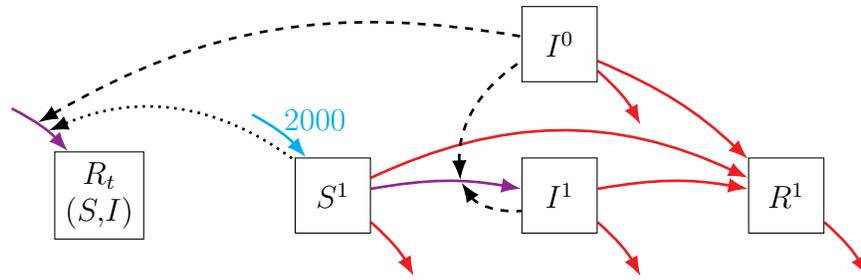


FIGURE A.8 – Modification du modèle de la figure A.7.

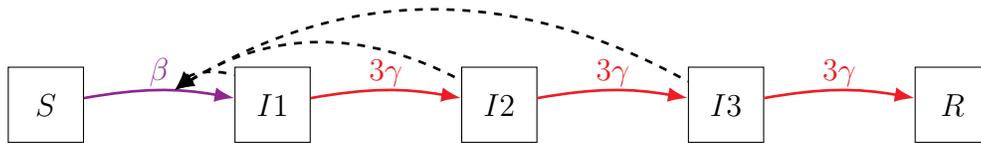


FIGURE A.9 – Modèle *SIR* avec séparation séquentielle des infectés.

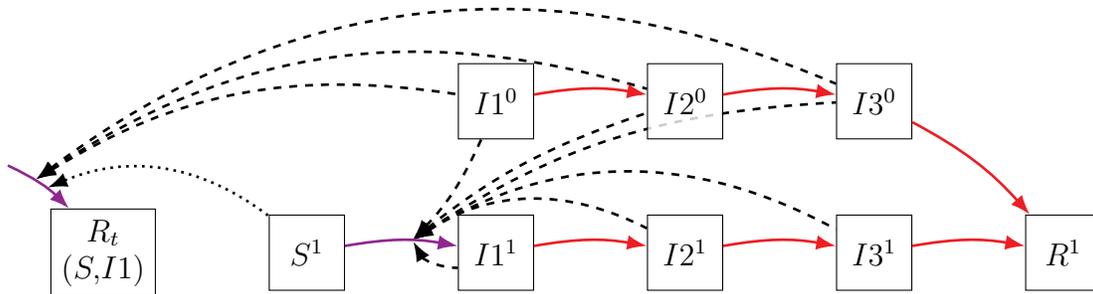


FIGURE A.10 – Modification du modèle de la figure A.9.

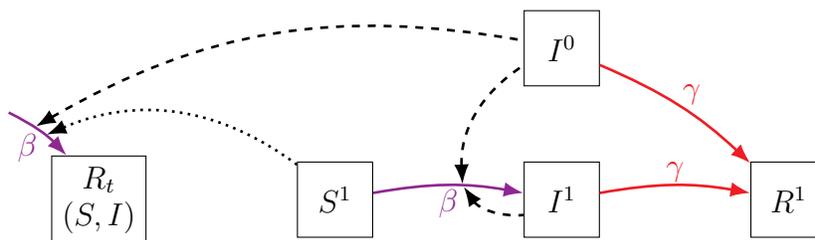


FIGURE A.11 – Modification du modèle *SIR* simple (figure 6.1).

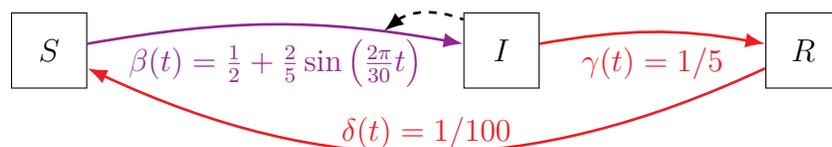


FIGURE A.12 – Modèle *SIRS* avec taux de contacts périodiques.



# Annexe B

---

## Quelques pseudocodes

La méthodologie présentée dans ce mémoire est implémentée dans une librairie écrite en *Python*. On présente dans cette annexe quelques exemples de fonctions pertinentes pour déterminer la chaîne d'infection d'un carquois (voir définition 3.1.1). Pour cette section,  $\{a : val(a), \dots\}$  représente un dictionnaire et  $[a, b, \dots]$  représente une liste ordonnée.

---

**Algorithme 1** Obtenir une liste contenant les nœuds de  $G$  à partir d'un dictionnaire  $G$ .

---

```
function V( $G$ )  
    return LIST( $G$ ["compartments"].KEYS())  
end function
```

---

---

**Algorithme 2** Obtenir un dictionnaire de tous les hyperarcs de  $G$  qui ne sont pas des *contacts*, sous la forme  $edges = \{u : A^-(u), v : A^-(v), \dots\}$ .

---

```
function EDGESCUT( $G$ )  
     $flows = G$ ["flows"] ▷ hyperarcs de  $G$ , en catégories  
     $compartments = V(G)$   
     $edges = \{\}$  for  $comp$  in  $compartments$  ▷ Initialiser dictionnaire vide  
  
    for  $category$  in  $flows$  do  
        for  $flow$  in  $category$  do  
             $to = flow$ ["to"]  
             $from = flow$ ["from"]  
             $contact = flow$ ["contact"]  
            if (not  $to$  in  $edges$ [ $from$ ]) and (not  $contact$ .STARTSWITH("Null")) then  
                 $edges$ [ $from$ ].APPEND( $to$ ) ▷ Ajouter les hyperarcs qui conviennent  
            end if  
        end for  
    end for  
    return  $edges$   
end function
```

---

---

**Algorithme 3** Trouver les chemins qui se trouvent entre *start* et *end* en respectant les hyperarcs donnés, avec les cycles. Inspiré de [11].

---

```

function SEARCHPATHS(start, end, visited, edges, path, allPaths)
  visited[start] = TRUE           ▷ visited prend en note les nœuds qui ont été visités
  path.APPEND(start)
  if start == end then           ▷ On a trouvé la fin du trajet
    allPaths.APPEND(path.COPY())
  else
    for descendant in edges[start] do           ▷ On descend dans le graphe
      if not visited[descendant] then           ▷ On recommence la recherche
        SEARCHPATHS(descendant, end, visited, edges, path, allPaths, cycles)
      else
        cycles.APPEND((path + [descendant]).COPY())
      end if
    end for
  end if
  path.POP()
  visited[start] = FALSE
  ▷ On a terminé avec cette recherche et on enlève donc le point de départ.
end function

```

---



---

**Algorithme 4** Trouver la chaîne d'infection de  $G$  qui se trouve entre *start* et *end*.

---

```

function SUBSET( $G$ , start, end)
  visited = {v : FALSE for v in  $V(G)$ }
  edges = EDGES CUT( $G$ )           ▷ On rejette les contacts
  allPaths = []                   ▷ Ce tableau sera mis à jour par SEARCHPATHS
  cycles = []                     ▷ Ce tableau sera également mis à jour
  SEARCHPATHS(start, end, visited, edges, [], allPaths, cycles)

  nodes = []
  for path in allPaths do           ▷ Union de tous les chemins
    nodes += [node for node in path if node not in nodes]
  end for
  for cycle in [cycle for cycle in cycles if cycle[-1] in nodes] do
    nodes += [node for node in cycle if node not in nodes]
  end for
  ▷ Ajout des cycles qui retournent dans la chaîne trouvée
end function

```

---

L'algorithme pour la modification de carquois lui-même est important pour ce projet de recherche, mais il sera omi de ce mémoire par souci de lisibilité. Cependant, le référentiel GitHub<sup>1</sup> sera rendu disponible lorsque ce mémoire sera soumis et il sera donc possible d'accéder à l'implémentation Python dans son entier.

1. <https://github.com/asimard1/AutomaticR0>