

# Mémoire du Magistère de Mathématiques : de l'exploration à l'approfondissement



Auriane GABAUT

Septembre 2020 - Octobre 2023

# Table des matières

,map size=0.9,widthh=

|                                                                             |           |
|-----------------------------------------------------------------------------|-----------|
| <b>Remerciements</b>                                                        | <b>8</b>  |
| <b>Introduction</b>                                                         | <b>9</b>  |
| <b>1 Cours au sein du Magistère de Mathématiques d'Orsay</b>                | <b>11</b> |
| 1.1 Première année : LDD3 . . . . .                                         | 12        |
| 1.2 Deuxième année : M1 MF . . . . .                                        | 13        |
| 1.3 Troisième année : M2 MSV . . . . .                                      | 14        |
| 1.4 Suite et perspective : Présentation d'un domaine de recherche . . . . . | 15        |
| Conclusion . . . . .                                                        | 16        |
| <b>I Exploration</b>                                                        | <b>17</b> |
| <b>2 Transcendance de <math>e</math> et <math>\pi</math></b>                | <b>18</b> |
| Introduction et Motivations . . . . .                                       | 18        |
| 2.1 Résultats Préliminaires . . . . .                                       | 20        |
| 2.1.1 Théorème fondamental des polynômes symétriques . . . . .              | 20        |

|           |                                                                                       |           |
|-----------|---------------------------------------------------------------------------------------|-----------|
| 2.1.2     | Quelques inégalités . . . . .                                                         | 25        |
| 2.2       | Transcendance de $e$ . . . . .                                                        | 27        |
| 2.2.1     | Résultat . . . . .                                                                    | 27        |
| 2.2.2     | Calcul de $J$ . . . . .                                                               | 28        |
| 2.2.3     | Encadrement de $ J $ . . . . .                                                        | 29        |
| 2.3       | Transcendance de $\pi$ . . . . .                                                      | 30        |
| 2.3.1     | Résultat . . . . .                                                                    | 30        |
| 2.3.2     | Autour de l'égalité d'Euler . . . . .                                                 | 31        |
| 2.3.3     | Calcul de $J$ . . . . .                                                               | 32        |
| 2.3.4     | Etude de $J$ . . . . .                                                                | 33        |
| 2.3.5     | Encadrement de $ J $ . . . . .                                                        | 36        |
| <b>3</b>  | <b>Rapport de stage de fin de licence</b>                                             | <b>37</b> |
|           | Introduction . . . . .                                                                | 37        |
| 3.1       | Première semaine du stage : phase d'observation . . . . .                             | 38        |
| 3.1.1     | Organisation de la classe . . . . .                                                   | 38        |
| 3.1.2     | Attendus . . . . .                                                                    | 38        |
| 3.2       | Tâches confiées . . . . .                                                             | 39        |
| 3.2.1     | Interrogations écrites . . . . .                                                      | 39        |
| 3.2.2     | DM et TD . . . . .                                                                    | 40        |
| 3.2.3     | Colles . . . . .                                                                      | 44        |
| 3.2.4     | Correction d'exercices et démonstrations . . . . .                                    | 48        |
|           | Conclusion . . . . .                                                                  | 55        |
| <b>II</b> | <b>Formation</b>                                                                      | <b>56</b> |
| <b>4</b>  | <b>Lars Hörmander et <i>The analysis of Linear Partial Differential Operators</i></b> | <b>57</b> |
|           | Introduction et Motivations . . . . .                                                 | 57        |
| 4.1       | Hörmander et ses travaux avant la médaille Fields . . . . .                           | 58        |
| 4.2       | Un apport conséquent dans le domaine des EDP . . . . .                                | 60        |

|          |                                                             |            |
|----------|-------------------------------------------------------------|------------|
| 4.3      | Reconnaissance et importance de son travail . . . . .       | 62         |
|          | Conclusion . . . . .                                        | 63         |
| <b>5</b> | <b>Introduction à l'analyse microlocale</b>                 | <b>65</b>  |
|          | Introduction et Motivations . . . . .                       | 66         |
| 5.1      | Rappels sur les distributions . . . . .                     | 66         |
| 5.1.1    | Définitions . . . . .                                       | 66         |
| 5.1.2    | Convergence de distributions . . . . .                      | 67         |
| 5.1.3    | Dérivation de distributions . . . . .                       | 67         |
| 5.1.4    | Produit de distributions . . . . .                          | 68         |
| 5.1.5    | Transformation de Fourier . . . . .                         | 68         |
| 5.2      | Vers le Front d'Onde . . . . .                              | 69         |
| 5.2.1    | Support et support singulier . . . . .                      | 69         |
| 5.2.2    | Le cône . . . . .                                           | 71         |
| 5.2.3    | Le front d'onde . . . . .                                   | 72         |
| 5.3      | Intégrales oscillantes . . . . .                            | 74         |
| 5.3.1    | Quelques résultats utiles . . . . .                         | 75         |
| 5.3.2    | Construction de l'intégrale oscillante . . . . .            | 78         |
| 5.3.3    | Front d'onde d'une intégrale oscillante . . . . .           | 81         |
| 5.4      | Du produit de distributions . . . . .                       | 83         |
| 5.4.1    | Généralités . . . . .                                       | 84         |
| 5.4.2    | Front d'onde et produit . . . . .                           | 86         |
| 5.5      | Front d'onde et équations aux dérivées partielles . . . . . | 99         |
| 5.5.1    | Théorème fondamental ( <i>Hörmander</i> ) . . . . .         | 100        |
| 5.5.2    | Cas d'un opérateur elliptique . . . . .                     | 103        |
| <b>6</b> | <b>Algorithme DQN</b>                                       | <b>105</b> |
|          | Introduction et Motivations . . . . .                       | 105        |
| 6.1      | Préliminaires . . . . .                                     | 107        |
| 6.1.1    | Bases de <i>Reinforcement Learning</i> . . . . .            | 107        |



|            |                                                               |            |
|------------|---------------------------------------------------------------|------------|
| 6.1.2      | $Q$ -Learning . . . . .                                       | 110        |
| 6.2        | Deep $Q$ -Learning . . . . .                                  | 111        |
| 6.2.1      | Le Deep $Q$ -Learning et les réseaux de neurones . . . . .    | 112        |
| 6.2.2      | Le $Q$ -Network . . . . .                                     | 113        |
| 6.2.3      | Mise en place de l'Algorithme DQN . . . . .                   | 114        |
| 6.2.4      | Double $Q$ -Learning . . . . .                                | 118        |
| 6.3        | Algorithme Rainbow . . . . .                                  | 122        |
| 6.3.1      | Prioritized Replay . . . . .                                  | 122        |
| 6.3.2      | Dueling Networks . . . . .                                    | 122        |
| 6.3.3      | Multi-Step Learning . . . . .                                 | 125        |
| 6.3.4      | Distributional RL . . . . .                                   | 125        |
| 6.3.5      | Noisy Nets . . . . .                                          | 126        |
| 6.3.6      | Rainbow . . . . .                                             | 128        |
| <b>III</b> | <b>Approfondissement</b>                                      | <b>130</b> |
| <b>7</b>   | <b>Estimation de l'efficacité contact Tracing</b>             | <b>131</b> |
|            | Introduction et Motivations . . . . .                         | 131        |
| 7.1        | Système Stochastique . . . . .                                | 132        |
| 7.1.1      | Le modèle . . . . .                                           | 132        |
| 7.1.2      | Modèle Mathématique . . . . .                                 | 133        |
| 7.1.3      | Modélisation pour une grande population . . . . .             | 135        |
| 7.2        | Approximate Bayesian Computation . . . . .                    | 140        |
| 7.2.1      | Principe Général . . . . .                                    | 140        |
| 7.2.2      | Construction de l'ABC . . . . .                               | 140        |
| 7.2.3      | Simulation . . . . .                                          | 141        |
| 7.3        | Estimation des paramètres pour des données simulées . . . . . | 142        |
|            | Conclusion . . . . .                                          | 145        |
| <b>8</b>   | <b>Procédure post-hoc</b>                                     | <b>146</b> |

|                                                                                               |            |
|-----------------------------------------------------------------------------------------------|------------|
| Introduction et Motivations . . . . .                                                         | 146        |
| 8.1 Analyse de données de protéomiques . . . . .                                              | 147        |
| 8.1.1 L'étude des protéomes . . . . .                                                         | 147        |
| 8.1.2 Analyse Protéomique par Spectrométrie de masse . . . . .                                | 148        |
| 8.1.3 Les outils statistiques usuels en protéomique . . . . .                                 | 149        |
| 8.2 Procédure Post-Hoc . . . . .                                                              | 150        |
| 8.2.1 Préliminaires . . . . .                                                                 | 150        |
| 8.2.2 Approche de Blanchard et al. . . . .                                                    | 157        |
| 8.2.3 Approche de Durand et al. . . . .                                                       | 159        |
| 8.2.4 Hybridation de bornes . . . . .                                                         | 161        |
| 8.3 Études de cas : applications des bornes Post-Hoc sur des données de protéomiques . . . .  | 162        |
| 8.3.1 Un premier jeu de données . . . . .                                                     | 162        |
| 8.3.2 Étude de cas sur un jeu de données contrôlées . . . . .                                 | 167        |
| Conclusion . . . . .                                                                          | 176        |
| <b>9 Stratégie de construction de modèle</b>                                                  | <b>178</b> |
| Introduction et Motivations . . . . .                                                         | 178        |
| 9.1 Préliminaires . . . . .                                                                   | 179        |
| 9.1.1 Modèles non-linéaire à effets mixtes . . . . .                                          | 180        |
| 9.1.2 Algorithmes de construction de modèle . . . . .                                         | 187        |
| 9.2 Sélection du modèle de covariables par LASSO dans SAMBA . . . . .                         | 191        |
| 9.2.1 Least Absolute Shrinkage and Selection Operator (LASSO) . . . . .                       | 191        |
| 9.2.2 Implémentation du LASSO pour la sélection de variable dans SAMBA . . . . .              | 193        |
| 9.2.3 Décroissance du critère de sélection au cours des itération . . . . .                   | 195        |
| 9.3 Application à la dynamique humorale post-boost . . . . .                                  | 196        |
| 9.3.1 Cadre d'application . . . . .                                                           | 196        |
| 9.3.2 Application à des simulations . . . . .                                                 | 198        |
| 9.3.3 Prevac-UP . . . . .                                                                     | 205        |
| <b>10 Utilisation de modèle à variables latentes pour la construction de modèle mécaniste</b> |            |

|                                                                                     |            |
|-------------------------------------------------------------------------------------|------------|
| <b>en grande dimension</b>                                                          | <b>210</b> |
| Introduction et Motivations . . . . .                                               | 211        |
| 10.1 Cadrage de la recherche . . . . .                                              | 212        |
| 10.1.1 Motivation . . . . .                                                         | 212        |
| 10.1.2 Modèles non linéaire à effets mixtes . . . . .                               | 213        |
| 10.1.3 Pertinence des modèles non-linéaires à effets mixtes en biologie . . . . .   | 215        |
| 10.1.4 Modèle à Variables Latentes . . . . .                                        | 215        |
| 10.1.5 L'exemple de la réponse immunitaire humorale . . . . .                       | 216        |
| 10.2 Analyse Méthodologique et Travail envisagé . . . . .                           | 218        |
| 10.2.1 Travail sur le modèle statistique . . . . .                                  | 218        |
| 10.2.2 Travail sur le modèle d'observation . . . . .                                | 219        |
| 10.2.3 L'étude Prevac-UP, et autres applications . . . . .                          | 220        |
| 10.3 Attendus . . . . .                                                             | 221        |
| 10.3.1 Calendrier des trois années de thèses . . . . .                              | 221        |
| 10.3.2 Perspectives . . . . .                                                       | 222        |
| <b>Annexe</b>                                                                       | <b>226</b> |
| A Cours en classe de MPSI . . . . .                                                 | 226        |
| B Algorithme DQN - Implémentation du jeu . . . . .                                  | 233        |
| C Algorithme de Construction de Modèle . . . . .                                    | 237        |
| C.1 Détails sur les algorithmes de sélection de covariables . . . . .               | 237        |
| C.2 Algorithme de validation croisée K-folds . . . . .                              | 239        |
| C.3 Paramètres individuels . . . . .                                                | 240        |
| C.4 Génération des covariables individuels . . . . .                                | 241        |
| C.5 Figures supplémentaires . . . . .                                               | 243        |
| C.6 Fits individuels . . . . .                                                      | 248        |
| D Présentation et soumission du travail lié au stage et en vu de la thèse . . . . . | 260        |
| D.1 Poster pour la journée des jeunes chercheurs du BPH . . . . .                   | 260        |
| D.2 Journée de Biostatistique - GDR de Toulouse . . . . .                           | 260        |

Merci !

,map size=0.9,widthh=

# Introduction

,map size=0.9,widthh=

Ce présent document se décompose en trois parties.

La première partie, que j'ai nommée "Exploration", concerne la première année de ce magistère. Cette section regroupe les moments de mon parcours qui, selon moi, m'ont aidé à faire le choix de ce que je voulais. J'ai d'abord pris des options, des cours, des sujets qui n'ont rien à voir avec ce que je fais actuellement, mais qui ont, à leur manière, contribué à m'y conduire. Cette partie contient notamment :

- Mon TER de L3 sur la transcendance de  $e$  et de  $\pi$ , encadré par Stéphane FISCHLER, avec mes camarades Antoine REMOVILLE et Ales BRAHITI ;
- Mon rapport de stage en classe préparatoire, au sein de la classe de MPSI2 du lycée Georges Clémenceau, à Reims, avec Pasquale DI FATTA.

La deuxième partie, nommée "Formation", correspond plutôt à la deuxième année et la première moitié de la troisième année du magistère, coïncidant avec la période où j'ai su précisément ce que je voulais faire. C'est à cette période que j'ai compris quelles étaient les mathématiques dont je souhaitais m'entourer dans le futur. Cette période a regroupé les premiers travaux encadrés en accord avec ce que je souhaitais faire plus tard, ils étaient plus alignés avec mon profil actuel mais encore diversifiés ; ils m'ont permis de passer de mon idée d'étudier dans un domaine large, comme l'analyse et l'apprentissage statistique, au domaine plus précis des biostatistiques et de la modélisation. Cette partie contient :

- Mon rapport écrit dans le cadre de l'UE d'histoire des sciences sur mon sujet de TER de M1 ;
- Mon sujet de TER de M1 sur l'analyse microlocale, encadré par Patrick GÉRARD et avec mes camarades Claire PLATEAUX et Enguérand TAMAGNA ;
- Mon projet pour l'UE de Reinforcement Learning au premier semestre de M2 ;

Enfin, la troisième et dernière partie "Approfondissement" représente ce que j'ai décidé de faire. Il s'agit de projets alignés avec mon profil actuel, plus récents, comme mon sujet de stage de M2 et

mon dossier de soumission de sujet de thèse. La voie finalement empruntée est celle de la modélisation pour les sciences médicales, notamment l'épidémiologie et la vaccinologie. L'idée est de travailler sur de la méthodologie, en profitant de l'apport théorique offert par le magistère, ajouté à mon M2 plutôt appliqué. Cette partie contient donc :

- Mon projet dans le cadre de l'UE de modélisation probabiliste et statistique pour l'épidémiologie au deuxième semestre de M2 ;
- Mon sujet de TER de M2 sur les procédures post-hoc pour les données de transcriptomiques, encadré par Marie CHION et Guillermo DURAND, avec mon camarade Alexandre PERRIN ;
- Mon rapport de stage de M2 sur la construction de modèles en grande dimension, dans l'équipe SISTM de l'Inria de Bordeaux, encadré par Mélanie PRAGUE ;
- Mon dossier de demande d'inscription en thèse sur le même thème, la construction de modèles en grande dimension, mais avec une nouvelle approche, encadré par Mélanie PRAGUE et Cécile PROUST-LIMA.

# Cours au sein du Magistère de Mathématiques d'Orsay



## Sommaire

|                                                                             |    |
|-----------------------------------------------------------------------------|----|
| 1.1 Première année : LDD3 . . . . .                                         | 12 |
| 1.2 Deuxième année : M1 MF . . . . .                                        | 13 |
| 1.3 Troisième année : M2 MSV . . . . .                                      | 14 |
| 1.4 Suite et perspective : Présentation d'un domaine de recherche . . . . . | 15 |
| Conclusion . . . . .                                                        | 16 |

Après 3 ans à étudier en classe préparatoire aux grandes écoles (MPSI/MP\*), j'ai décidé de poursuivre mes études en magistère de mathématiques. Mes débuts avec le magistère n'ont pas été évidents, et bien que mon idée sur ma future carrière était fixée depuis longtemps, mes décisions et projets d'études n'ont pas suivi une route droite et déterminée.

Quand j'étais plus jeune, aussi loin que je puisse me rappeler et jusqu'à mes 16 ans, je souhaitais devenir vétérinaire, et seulement en deuxième position, m'orienter vers les mathématiques ou la médecine. En entrant en filière scientifique au lycée, mon attrait pour les mathématiques a grandi, et c'est finalement vers des études de mathématiques que je me suis dirigée. Ma première idée, en sortant du lycée et en entrant en C.P.G.E., était de devenir professeur de mathématiques dans l'enseignement supérieur ou enseignante-chercheuse ; je pensais donc faire ma prépa pour acquérir une solide formation, puis aller à l'université pour préparer l'agrégation et éventuellement effectuer une thèse. Finalement, j'ai entendu parler des magistères et j'ai donc voulu intégrer celui d'Orsay, ce que j'ai eu la chance de réaliser en septembre 2020.

J'ai donc finalement passé mes 3 années d'études suivantes à Orsay, en tâtonnant d'abord sur les domaines des mathématiques m'intéressant le plus, mais tout en gardant en tête mon désir de m'orienter vers la recherche et/ou l'enseignement. Le magistère m'a finalement permis de concilier mes aspirations antérieures en associant les mathématiques que j'avais apprises et aimées au domaine médical qui m'attirait.

---

## 1.1. Première année : LDD3

---

Comme évoqué précédemment, cette première année de magistère m'a plutôt permise d'explorer les différents domaines des mathématiques. À l'époque, je n'avais aucune idée claire de ce que je voulais faire en matière de master, ni même quel domaine des mathématiques pourrait m'intéresser pour toute ma future carrière.

Avec le recul, je considère que la troisième année de licence et la première année de magistère proposées à Orsay étaient parfaites dans une situation comme la mienne. Nous avions un tronc commun très diversifié, comprenant des cours d'algèbre, de calcul différentiel, de topologie, mais aussi de théorie de la mesure et de probabilité. En ce qui concerne les options, lors de mon année, nous avions la possibilité de choisir au premier semestre entre deux options : combinatoire algébrique et algorithmie, et j'ai opté pour l'algorithmie. Au deuxième semestre, il était nécessaire de choisir au moins une unité d'enseignement (UE) pour une MAO, Mathématiques Assistées par Ordinateur, parmi Calcul Formel, Probabilités et Statistiques, et Calcul Scientifique, et j'ai choisi Calcul Formel. De plus, au même semestre, il y avait une option à choisir entre Informatique théorique et Physique-Mathématique, ou bien suivre les deux autres UE de MAO restantes. J'ai décidé de suivre le cours de Physique-Mathématique.

Mes choix d'options pour cette année avaient été faits principalement par curiosité envers l'une ou l'autre des options, en tenant compte du fait que mes résultats dans des matières telles que l'algèbre ou le calcul différentiel étaient meilleurs. Cependant, au cours de cette année, j'ai réalisé que l'analyse était un domaine qui m'intéressait plus largement. À présent, il ne semble pas pertinent d'avoir suivi l'UE de Calcul Formel, par exemple, alors que je me suis dirigée exclusivement vers l'analyse, et plus précisément vers les statistiques et les probabilités. Cependant, ma curiosité avait primé sur l'intérêt, et je pense que ces choix d'options m'ont permis de comprendre ce que je voulais réellement faire. Au cours de cette même année, j'ai effectué mon TER, présenté au chapitre 2, supervisé par Stéphane FISCHLER, portant sur la transcendance de  $e$  et de  $\pi$ .

À la fin de ma L3, en raison du COVID, le stage était facultatif et le contexte sanitaire a rendu difficile la recherche de stage. Malgré ma volonté d'en effectuer un, je n'ai pas eu la chance d'en trouver un à la hauteur de mes espérances en laboratoire, avec la contrainte de ne pas être un laboratoire de mathématiques exclusivement. J'ai finalement effectué un stage en classe préparatoire, sur autorisation des responsables de formation, dans l'idée de confirmer ou d'infirmer mon envie de faire de l'enseignement. Le rapport relatif à ce dernier est présenté au chapitre 3. À la fin de ma L3, j'ai décidé de poursuivre en master de Mathématiques Fondamentales, envisageant de me spécialiser davantage dans le domaine de l'analyse.



---

## 1.2. Deuxième année : M1 MF

---

Pour ma deuxième année de magistère et première année de master, j'ai fait ma rentrée en master de Mathématiques Fondamentales, en septembre 2021, toujours à Orsay. Comme expliqué, c'est vers l'analyse que je me suis dirigée d'abord, et dès la présentation de rentrée des masters, j'ai décidé quel M2 je souhaitais : aller en MSV, le master de Mathématiques pour les Sciences du Vivant. Lors de cette même présentation en L3, le choix du master MSV ne s'était pas imposé. Peut-être par manque d'attention, ou parce que l'idée n'avait pas encore raisonné en moi comme à ce moment-là. Le master MSV proposé à Paris-Saclay s'est donc imposé comme mon objectif pour cette année : pour la prédominance de manière analytique -que ce soit des équations différentielles, des statistiques, ou de l'apprentissage statistique, les trois sous-domaines que j'avais alors décidé de cibler- ou pour la possibilité de travailler pour des enjeux médicaux et sanitaires.

Le choix de mes cours en master a donc été complètement orienté en fonction des prérequis de MSV. Au premier semestre, mes matières principales étaient donc Distribution et Analyse de Fourier par Stéphane NONNENMACHER, un cours que j'ai particulièrement adoré, Probabilités par Nicolas CURIEN, avec l'option Mathématiques pour l'Intelligence Artificielle 1 par Gilles BLANCHARD en cours magistère, UE difficile pour moi à l'époque en raison du manque de base en statistiques. Au deuxième semestre, j'ai suivi le cours passionnant d'Équations aux Dérivées Partielles d'Evolution par Matthieu LÉAUTAUD, le cours de Statistiques par Élisabeth GASSIAT et pour le magistère, le cours d'Optimisation à l'ENS par Alain TROUVÉ. Enfin, pour la dernière partie de l'année, j'ai suivi l'UE de MAO Probabilité et Statistiques, ainsi que l'UE d'Histoire des Mathématiques, en complément, dont le rapport est présenté au chapitre 4. Pour cette année, j'ai effectué mon TER sur une introduction à l'Analyse Microlocale, dont le mémoire est au chapitre 5, supervisé par Patrick GÉRARD.

Comme je l'espérais, cette année a été de très loin l'une de mes meilleures années d'études. Les cours que j'ai consciencieusement choisis m'ont passionnée, motivée plus que je ne l'espérais, et m'ont grandement aidée à reprendre des habitudes étudiantes plus normales et épanouissantes. À l'issue de cette année, j'ai intégré comme je le souhaitais le M2 MSV et ai aussi obtenu la bourse Sophie Germain de la FMJH, fondation de mathématiques Jacques Hadamard. Cette année-là, mon responsable de master et encadrant de TER, Patrick GÉRARD, a joué un rôle crucial en me poussant à candidater pour cette bourse, en me guidant dans mes choix de cours, et en m'accompagnant sur des démarches liées à ma situation, notamment avec les assistantes sociales du CROUS. J'envisageais néanmoins à l'époque d'effectuer la césure de magistère pour passer l'agrégation, mais ce choix a été écarté face à l'opportunité fournie par la bourse de m'offrir une année de M2 sereine, centrée exclusivement sur mes études.

---

### 1.3. Troisième année : M2 MSV

---

En septembre 2022, j'ai donc entamé ma deuxième année de master en Mathématiques pour les Sciences du Vivant. Malheureusement, le début de ce master a été assez décevant. Le tronc commun du premier semestre ne correspondait pas exactement à mes attentes. Pour commencer, nous avons eu une UE de biologie qui n'a duré qu'une semaine, et dont l'examen consistait en une lecture de trois articles suivie de quelques questions de compréhension. Cette UE m'a grandement déçue, car son contenu m'a semblé très superficiel. En ce qui concerne les UE de mathématiques, il y avait une UE de Modélisation Déterministe, qui ressemblait beaucoup à celle que j'avais suivie en L3. Il y avait aussi une UE d'Optimisation et de Simulation Numérique, qui était certes intéressante et bien présentée, mais qui ressemblait également beaucoup à celle que j'avais suivie en M1. Heureusement, nous avons eu une UE passionnante sur les Processus Stochastiques enseignée par Sylvie MÉLÉARD. De plus, nous avons le choix entre deux options : Statistique en Grande Dimension avec Christophe GIRAUD et Apprentissage par Renforcement avec Erwan LE PENNEC. J'ai opté pour l'UE d'Apprentissage par Renforcement, et j'ai particulièrement apprécié le projet que j'ai réalisé en fin de semestre. Le rapport de ce projet est présenté dans le chapitre 6.

Au deuxième semestre, nous avons eu la liberté de choisir toutes nos UE, ce qui m'a permis de tirer plus profit du master MSV. J'ai décidé de suivre des UE liées à la modélisation statistique, probabiliste et déterministe afin de maintenir une diversité dans mes compétences. J'ai donc choisi de suivre l'UE de Méthodes Statistiques en Grande Dimension pour l'Analyse de Données "-omiques" par Céline LÉVY-LEDUC, l'UE de Modélisations Probabilistes et Statistiques pour l'Epidémiologie par Viet Chí TRAN, dont le rapport de projet relatif à l'UE de Viet Chí TRAN est présenté dans le chapitre 7; et l'UE de Modèles à Variables Latentes en Biologie et Ecologie par Sophie DONNET. J'ai également assisté, et impulsé l'organisation, d'un mini-cours sur les Modèles Non-Linéaires à Effets Mixtes enseigné par Mélanie PRAGUE pendant les vacances de février. En tant qu'auditrice libre, j'ai suivi l'UE de Sélection de Modèle par Pascal MASSART, qui était proposée dans le cadre des Mathématiques de l'Aléatoire (MDA), ainsi que l'UE de Causalité et Données Manquantes par Marianne CLAUDEL, dispensée dans le cadre du master Data Science (MDS). Toutes ces UE se sont révélées passionnantes et parfaitement en adéquation avec mes objectifs d'études. Au cours de ce second semestre, j'ai eu l'opportunité d'enrichir considérablement ma formation.

Tout au long de ma deuxième année de master, j'ai consacré mon TER à l'étude des procédures post-hoc pour l'analyse de données de transcriptomique, sous la supervision de Guillermo DURAND et Marie CHION. Ce TER m'a semblé être ma première véritable expérience de recherche, avec une question ouverte à explorer, similaire à un stage de master. Le mémoire résultant de ce travail est présenté dans le chapitre 8. Par la suite, j'ai effectué mon stage du 1er avril 2023 au 30 septembre 2023 au sein de l'équipe SISTM de l'Inria de Bordeaux, supervisé par Mélanie PRAGUE. Ce stage portait sur les stratégies de construction de modèles en grande dimension. Le mémoire de stage est présenté dans le chapitre 9. Désormais, je souhaite me consacrer entièrement à la modélisation mécaniste en grande dimension, c'est-à-dire à l'utilisation de modèles statistiques basés sur des Équations Différentielles Ordinaires (EDO) ou des Équations aux Dérivées Partielles (EDP), avec un grand nombre de prédictors, pour étudier des

phénomènes observés chez un petit nombre d'individus. Suite à ce stage, j'ai entamé une thèse portant sur l'utilisation de modèles à variables latentes dans le but d'améliorer les modèles mécanistes en grande dimension, dont les détails sont données ensuite.

---

## 1.4. Suite et perspective : Présentation d'un domaine de recherche

---

En biologie, notamment en vaccinologie et en épidémiologie, la compréhension des phénomènes biologiques, parfois très complexes sur le plan mécaniste, est d'une importance cruciale pour le développement de stratégies médicales efficaces. Il est essentiel d'utiliser les modèles statistiques les plus pertinents possible, en tenant compte des connaissances mécanistes et biologiques préalables, ainsi que des données collectées lors des essais cliniques.

L'une des problématiques majeures dans ce domaine de recherche concerne la sélection des caractéristiques ou covariables d'un individu qui influencent le processus. Par exemple, le fait que l'on soit un homme, une femme, majeur ou mineur, etc, peut influencer sur la réponse à un vaccin, un médicament ou une maladie. Par conséquent, il est impératif de pouvoir identifier les informations significatives en vue d'affiner les stratégies médicales et de personnaliser au maximum les traitements.

Cette problématique se pose d'abord, comme introduit, à l'échelle des données démographiques, telles que le sexe, l'âge, l'origine ethnique, etc. Pour un petit ensemble de données, nous pouvons construire algorithmiquement le modèle final le plus pertinent, en utilisant les données disponibles et les connaissances préalables. Cependant, lorsqu'il s'agit de données génétiques et de la détermination des gènes jouant un rôle dans un phénomène biologique, la construction d'un modèle devient un véritable défi statistique. Ces données, qualifiées de "grande dimension", car le nombre de prédicteurs potentiels (les gènes) dépasse largement le nombre d'individus observés, nécessitent de nouveaux outils d'analyse.

Il existe déjà des méthodes statistiques permettant d'incorporer ces données de grande dimension dans des cadres simples tels que les modèles linéaires. Cependant, si l'on considère maintenant le contexte des modèles non linéaires à effets mixtes, l'intégration de données de grande dimension devient un véritable défi. C'est précisément le sujet de mon stage, détaillé davantage dans le chapitre 9, mais aussi le thème de recherche de ma thèse, présenté en chapitre 10.

---

## Conclusion

---

Le choix de poursuivre mes études en magistère de mathématiques a été une décision cruciale dans mon parcours académique. À travers ces années de formation, j'ai pu explorer divers domaines des mathématiques, affiner mes intérêts, et finalement trouver ma voie dans le monde académique. Retourner en arrière et analyser l'ensemble de mon parcours, c'est comprendre l'intérêt fondamental que le magistère a eu pour moi.

Lorsque j'ai intégré le magistère, mes aspirations professionnelles étaient encore floues. Mais grâce à mes années d'étude au magistère, j'ai pu combiner des aspirations divergentes en explorant des sujets de mathématiques liés à la médecine et aux sciences du vivant. Le large éventail de cours proposés et la flexibilité offerte par le magistère m'a permis de personnaliser mon parcours en fonction de mes intérêts.

Une autre facette cruciale du magistère a été la possibilité d'interagir avec des enseignants et des chercheurs de renom qui ont guidé et encouragé mon développement académique. Leur soutien a été essentiel pour ma progression et m'a ouvert des portes vers des opportunités que je n'aurais pas su saisir seule.

En conclusion, le magistère de mathématiques a été un véritable tremplin dans mon parcours académique et professionnel. Il m'a aidée à clarifier mes aspirations, à développer mes compétences, et à forger mon identité en tant que future mathématicienne -du moins, je l'espère.

# Première Partie : Exploration

,max size=0.9,widthh=

# Transcendance de $e$ et $\pi$



BRAHITI Ales, GABAUT Auriane, REMOVILLE Antoine  
Sous la direction de S. FISCHLER

Avril 2021

## Sommaire

|                                                      |           |
|------------------------------------------------------|-----------|
| <b>Introduction et Motivations</b>                   | <b>18</b> |
| <b>2.1 Résultats Préliminaires</b>                   | <b>20</b> |
| 2.1.1 Théorème fondamental des polynômes symétriques | 20        |
| 2.1.2 Quelques inégalités                            | 25        |
| <b>2.2 Transcendance de <math>e</math></b>           | <b>27</b> |
| 2.2.1 Résultat                                       | 27        |
| 2.2.2 Calcul de $J$                                  | 28        |
| 2.2.3 Encadrement de $ J $                           | 29        |
| <b>2.3 Transcendance de <math>\pi</math></b>         | <b>30</b> |
| 2.3.1 Résultat                                       | 30        |
| 2.3.2 Autour de l'égalité d'Euler                    | 31        |
| 2.3.3 Calcul de $J$                                  | 32        |
| 2.3.4 Etude de $J$                                   | 33        |
| 2.3.5 Encadrement de $ J $                           | 36        |

## Introduction et Motivations

Notre travail a pour but de prouver la transcendance de deux constantes centrales en mathématiques,  $e$  et  $\pi$ . Pour ce faire, nous nous appuyerons sur un extrait du livre d'Alan BAKER [1].

Nombre algébrique, nombre transcendant

Définition -

Un complexe est dit algébrique s'il existe un polynôme non nul à coefficients entiers qui l'annule. Un complexe est dit transcendant s'il n'est pas algébrique.

Remarquons que, pour  $\alpha$  un complexe, on a l'équivalence suivante :

$$\exists P \in \mathbb{Q}[X], P \neq 0 \text{ tel que } P(\alpha) = 0 \iff \exists P \in \mathbb{Z}[X], P \neq 0 \text{ tel que } P(\alpha) = 0$$

Le sens réciproque est direct comme  $\mathbb{Z} \subset \mathbb{Q}$ . Quant au sens direct, si l'on note  $P = \sum_{i=0}^d \frac{p_i}{q_i} X^i$  un polynôme non nul annulateur de  $\alpha$ , il suffit de remarquer que  $q_0 \cdots q_d P$  est un polynôme non nul annulateur de  $\alpha$  à coefficients entiers.

Pour  $\alpha$  un nombre algébrique, il est intéressant de noter que l'ensemble des polynômes qui annulent  $\alpha$  est un idéal de  $\mathbb{Q}[X]$ , qui est principal. Dès lors, il existe un unique polynôme unitaire qui l'engendre, ce qui justifie la définition suivante.

*Définition -*

Si  $\alpha$  est un nombre algébrique, on appelle polynôme minimal de  $\alpha$  le polynôme unitaire non nul de degré minimal qui annule  $\alpha$ . On appelle degré de  $\alpha$  le degré du polynôme minimal associé.

### *Un peu d'histoire autour des nombres transcendants*

La question de l'existence de nombres transcendants aurait émergé au XVII<sup>ème</sup> siècle et Gottfried Wilhelm LEIBNIZ serait le premier à s'être demandé si de tels nombres existaient. Un siècle plus tard, Leonhard EULER démontre en 1744 l'irrationalité de  $e$  et Jean-Henri LAMBERT celle de  $\pi$  en 1761. En 1873, Charles HERMITE propose la première preuve de la transcendance de  $e$  et Ferdinand VON LINDERMANN fournit la première preuve de la transcendance de  $\pi$  en 1882.

Depuis, les nombres transcendants ont continué d'intéresser les mathématiciens. Le septième problème de Hilbert, énoncé en 1900, propose de démontrer le résultat suivant : si  $\alpha$  est un nombre algébrique différent de 0 et de 1 et que  $\beta$  est un nombre algébrique irrationnel alors  $\alpha^\beta$  est transcendant. Quelques années plus tard, Aleksandr GELFOND et Theodor SCHNEIDER prouvent ce résultat, qui fournit des exemples de nombres transcendants, et donnèrent alors leur nom au nombre transcendant  $2^{\sqrt{2}}$ , dorénavant appelé la constante de Gelfond-Schneider.

### *Quelques exemples de nombres transcendants*

Comme annoncé, nous tâcherons de montrer la transcendance de  $e$  et de  $\pi$ , qui constitue un premier exemple de nombres transcendants. Il en existe d'autres dont le caractère transcendant a été prouvé : la constante de Gelfond-Schneider  $2^{\sqrt{2}}$ ,  $\sin(1)$  ou encore  $e^\pi$ .

### *La transcendance de $\pi$ et le problème de la quadrature du cercle*

De l'antiquité grecque jusqu'au XIX<sup>e</sup> siècle, le problème de la quadrature du cercle faisait partie de trois importants problèmes de géométrie, avec la duplication du cube et la trisection de l'angle. L'objectif était de construire un carré d'aire égale à celle d'un cercle donné.

La quadrature d'un cercle de rayon  $r$  consiste à construire un carré d'aire  $\pi r^2$ . De manière équivalente, il faudrait donc réussir à construire à la règle et au compas  $\sqrt{\pi}$ . On dira alors que  $\sqrt{\pi}$  est constructible.

En 1837, Pierre-Laurent WANTZEL prouve que les deux derniers problèmes évoqués sont impossibles et démontre un théorème - le théorème de Wantzel- duquel il déduit une condition nécessaire pour qu'un nombre soit constructible : si un nombre est constructible, alors il est algébrique. Par la suite, en 1844, Joseph LIOUVILLE met en évidence l'existence des nombres transcendants, et c'est en 1882 que Ferdinand VON LINDERMANN, en démontrant la transcendance de  $\pi$ , prouve l'impossibilité de la quadrature du cercle.

---

## 2.1. Résultats Préliminaires

---

### 2.1.1. Théorème fondamental des polynômes symétriques

#### Résultat et Lemmes préparatoires

On considère  $A$  un anneau commutatif, nous allons nous intéresser à l'anneau des polynômes sur  $A$  à  $n$  indéterminées noté  $A[X_1, \dots, X_n]$ . Dans le but de démontrer la transcendance de  $\pi$ , nous utiliserons le théorème fondamental des polynômes symétriques. Pour l'énoncer et le démontrer, nous nous appuierons sur une leçon pour l'agrégation [2] ainsi que sur un extrait du livre de X. GOURDON [3].

#### Définition 2.1.1 -

Pour  $q \in \llbracket 1; n \rrbracket$ , on définit  $\Sigma_q$  le  $q^{\text{e}}$  polynôme symétrique élémentaire de  $A[X_1, \dots, X_n]$  :

$$\Sigma_q = \sum_{1 \leq i_1 < \dots < i_q \leq n} X_{i_1} \dots X_{i_q}$$

#### Définition 2.1.2 -

Un polynôme  $P$  de  $A[X_1, \dots, X_n]$  est symétrique si, pour toute permutation  $\sigma$  de  $\llbracket 1; n \rrbracket$ , on a :

$$P(X_{\sigma(1)}, \dots, X_{\sigma(n)}) = P(X_1, \dots, X_n)$$

#### Théorème 2.1.1 -

Soient  $n \in \mathbb{N}^*$  et  $P \in A[X_1, \dots, X_n]$  où  $A$  est un anneau commutatif intègre. Si  $P$  est symétrique alors il existe un unique polynôme  $Q \in A[\Sigma_1, \dots, \Sigma_n]$  tel que :

$$P(X_1, \dots, X_n) = Q(\Sigma_1, \dots, \Sigma_n)$$

Deux lemmes sont nécessaires pour démontrer ce résultat :

#### Lemme 2.1.1 -

Pour  $q \in \llbracket 1; n-1 \rrbracket$ , en notant  $(\Sigma_q)_0$  le polynôme  $\Sigma_q$  évalué en  $X_n = 0$ ,  $(\Sigma_q)_0$  est le  $q^{\text{e}}$  polynôme symétrique élémentaire de  $A[X_1, \dots, X_{n-1}]$ .

*Démonstration.* Notons  $\Sigma'_q$  le  $q^{\text{e}}$  polynôme symétrique élémentaire de  $A[X_1, \dots, X_{n-1}]$



On a :

$$\begin{aligned}
 \Sigma_q &= \sum_{1 \leq i_1 < \dots < i_q \leq n} X_{i_1} \dots X_{i_q} \\
 &= \sum_{\substack{1 \leq i_1 < \dots < i_q \leq n \\ i_q \neq n}} X_{i_1} \dots X_{i_q} + \sum_{\substack{1 \leq i_1 < \dots < i_q \leq n \\ i_q = n}} X_{i_1} \dots X_{i_q} \\
 &= \sum_{1 \leq i_1 < \dots < i_q \leq n-1} X_{i_1} \dots X_{i_q} + \sum_{1 \leq i_1 < \dots < i_{q-1} \leq n-1} X_{i_1} \dots X_{i_{q-1}} X_n
 \end{aligned}$$

d'où, pour  $X_n = 0$  :

$$(\Sigma_q)_0 = \sum_{1 \leq i_1 < \dots < i_q \leq n-1} X_{i_1} \dots X_{i_q} = \Sigma'_q$$

□

Lemme 2.1.1 -

Soit  $P \in A[X_1, \dots, X_n]$  un polynôme symétrique.  
Si  $P(X_1, \dots, X_{n-1}, 0) = 0$ , alors  $X_1 \dots X_n | P$ .

*Démonstration.* Soit  $P \in A[X_1, \dots, X_n]$  un polynôme symétrique.

Supposons que  $P(X_1, \dots, X_{n-1}, 0) = 0$ . Notons  $P = \sum_{i \in \mathbb{N}^n} a_i X_1^{i_1} \dots X_n^{i_n}$ .

avec  $(a_i)_{i \in \mathbb{N}^n} \in A^{\mathbb{N}^n}$  nulle à partir d'un certain rang, et  $i = (i_1, \dots, i_n) \in \mathbb{N}^n$ .

$$P(X_1, \dots, X_{n-1}, 0) = \sum_{\substack{i \in \mathbb{N}^n \\ i_n = 0}} a_i X_1^{i_1} \dots X_{n-1}^{i_{n-1}} = 0$$

On a donc  $a_i = 0$ , pour tout  $i \in \mathbb{N}^n$  tel que  $i_n = 0$ .

$$\begin{aligned}
 P(X_1, \dots, X_n) &= \sum_{\substack{i \in \mathbb{N}^n \\ i_n \neq 0}} a_i X_1^{i_1} \dots X_n^{i_n} \\
 &= X_n \sum_{\substack{i \in \mathbb{N}^n \\ i_n \geq 1}} a_i X_1^{i_1} \dots X_n^{i_n-1}
 \end{aligned}$$

Donc  $X_n | P$ , et  $P$  étant symétrique, on conclut que  $X_k | P$ , pour tout  $k$ . On a alors que  $X_1 \dots X_n | P$ . □

Revenons maintenant à la démonstration du théorème.

*Existence du polynôme  $Q$*

Nous allons procéder par double récurrence pour montrer, pour tout polynôme  $P$  symétrique dans  $A[X_1, \dots, X_n]$ , l'existence d'un polynôme  $Q$  tel que :

$$P(X_1, \dots, X_n) = Q(\Sigma_1, \dots, \Sigma_n)$$

Soit  $\mathcal{A}_n$  correspondant à la proposition : "Pour tout polynôme  $P \in A[X_1, \dots, X_n]$  symétrique, il existe un polynôme  $Q \in A[X_1, \dots, X_n]$  tel que  $P(X_1, \dots, X_n) = Q(\Sigma_1, \dots, \Sigma_n)$ ".

On montre par récurrence sur  $n \in \mathbb{N}^*$  la proposition  $\mathcal{A}_n$ .

Pour  $n = 1$  : On se place dans  $A[X_1]$ , on obtient alors  $\Sigma_1 = X_1$ .

En prenant  $Q(X_1) = P(X_1)$ , on a bien  $Q \in A[X_1]$  et  $P(X_1) = Q(\Sigma_1)$ .

La proposition est donc bien vérifiée pour  $n = 1$ .

On suppose établie la proposition pour un certain rang  $n - 1 \in \mathbb{N}^*$  :

Soit  $\mathcal{A}_n(p)$  la proposition : "l'existence dans le théorème est vérifiée pour tout polynôme de degré total  $p \in \mathbb{N}^*$  à  $n \in \mathbb{N}^*$  indéterminées".

— Pour  $p = 0$  :

Pour un polynôme constant  $P$ , on pose  $Q = P$  :

On a bien  $Q \in A[X_1, \dots, X_n]$  et  $P(X_1, \dots, X_n) = c = Q(\Sigma_1, \dots, \Sigma_n)$ .

— On suppose  $\mathcal{A}_n(p)$  établie jusqu'à un certain rang  $p - 1 \in \mathbb{N}^*$  :

Soit  $P \in A[X_1, \dots, X_n]$  symétrique de degré  $p$ . Notons  $T$  le polynôme tel que  $T(X_1, \dots, X_{n-1}) = P(X_1, \dots, X_{n-1}, 0)$ .

$T \in A[X_1, \dots, X_{n-1}]$  est symétrique car  $P$  l'est : en effet, pour tout  $\sigma \in \mathcal{S}_{n-1}$ , on a :

$$T(X_{\sigma(1)}, \dots, X_{\sigma(n-1)}) = P(X_{\sigma(1)}, \dots, X_{\sigma(n-1)}, 0)$$

Or en posant  $\tilde{\sigma}(i) = \sigma(i)$  si  $1 \leq i < n$  et  $\tilde{\sigma}(n) = n$ , on a bien  $\tilde{\sigma} \in \mathcal{S}_n$ . Et :

$$\begin{aligned} P(X_{\tilde{\sigma}(1)}, \dots, X_{\tilde{\sigma}(n)}) &= P(X_{\tilde{\sigma}(1)}, \dots, X_{\tilde{\sigma}(n-1)}, X_n) \\ \underbrace{P(X_1, \dots, X_n)}_{\text{car } P \text{ est symétrique}} &= P(X_{\sigma(1)}, \dots, X_{\sigma(n-1)}, X_n) \end{aligned}$$

D'où :

$$P(X_{\sigma(1)}, \dots, X_{\sigma(n-1)}, 0) = P(X_1, \dots, X_{n-1}, 0)$$

Et finalement :

$$T(X_{\sigma(1)}, \dots, X_{\sigma(n-1)}) = P(X_1, \dots, X_{n-1}, 0) = T(X_1, \dots, X_{n-1})$$

$T$  est donc un polynôme symétrique de  $A[X_1, \dots, X_{n-1}]$  ; on peut appliquer l'hypothèse de récurrence  $\mathcal{A}_{n-1}$  :

$$\exists Q_1 \in A[X_1, \dots, X_{n-1}], \quad T(X_1, \dots, X_{n-1}) = Q_1(\Sigma'_1, \dots, \Sigma'_{n-1})$$

où  $\Sigma'_q$  est le  $q^{\text{e}}$  polynôme symétrique élémentaire de  $A[X_1, \dots, X_{n-1}]$ .

D'après le Lemme 1 :

$$T(X_1, \dots, X_{n-1}) = Q_1((\Sigma_1)_0, \dots, (\Sigma_{n-1})_0)$$

Posons à présent :

$$P_1(X_1, \dots, X_n) = P(X_1, \dots, X_n) - Q_1(\Sigma_1, \dots, \Sigma_{n-1}) \quad (2.1)$$

$P_1 \in A[X_1, \dots, X_n]$  est symétrique : en effet, pour  $\sigma \in \mathcal{S}_n$ , on a :

$$\begin{aligned}\Sigma_q^\sigma &= \sum_{1 \leq i_1 < \dots < i_q \leq n} X_{\sigma(i_1)} \dots X_{\sigma(i_q)} \\ &= \sum_{1 \leq j_1 < \dots < j_q \leq n} X_{j_1} \dots X_{j_q} \\ &= \Sigma_q\end{aligned}$$

Ainsi :

$$\begin{aligned}P_1(X_{\sigma(1)}, \dots, X_{\sigma(n)}) &= P(X_{\sigma(1)}, \dots, X_{\sigma(n)}) - Q_1(\Sigma_1^\sigma, \dots, \Sigma_{n-1}^\sigma) \\ &= P(X_1, \dots, X_n) - Q_1(\Sigma_1, \dots, \Sigma_{n-1}) \\ &= P_1(X_1, \dots, X_n)\end{aligned}$$

De plus,  $\deg(P) = p$  et  $\deg(Q_1) = \deg(T) \leq p$ , donc  $\deg(P_1) \leq p$ .

En évaluant  $P_1$  pour  $X_n = 0$ , on observe que :

$$P_1(X_1, \dots, X_{n-1}, 0) = T(X_1, \dots, X_{n-1}) - Q_1((\Sigma_1)_0, \dots, (\Sigma_{n-1})_0) = 0$$

Donc d'après le Lemme 2 :

$$\underbrace{X_1 \dots X_n}_{\Sigma_n} | P_1$$

Donc il existe un polynôme  $P_2 \in A[X_1, \dots, X_n]$  de degré  $k \leq p - n$ , tel que  $P_1 = \Sigma_n P_2$ .

$P_2$  est symétrique : en effet, pour  $\sigma \in \mathcal{S}_n$  :

$$P_1(X_{\sigma(1)}, \dots, X_{\sigma(n)}) = P_1(X_1, \dots, X_n)$$

D'où :

$$\begin{aligned}X_1 \dots X_n P_2(X_1, \dots, X_n) &= X_{\sigma(1)} \dots X_{\sigma(n)} P_2(X_{\sigma(1)}, \dots, X_{\sigma(n)}) \\ &= X_1 \dots X_n P_2(X_{\sigma(1)}, \dots, X_{\sigma(n)})\end{aligned}$$

Donc :

$$\underbrace{\Sigma_n}_{\neq 0} (P_2(X_1, \dots, X_n) - P_2(X_{\sigma(1)}, \dots, X_{\sigma(n)})) = 0$$

Or comme  $A$  est intègre,  $A[X_1, \dots, X_n]$  l'est aussi, et on a donc :

$$P_2(X_1, \dots, X_n) = P_2(X_{\sigma(1)}, \dots, X_{\sigma(n)})$$

On peut donc appliquer l'hypothèse de récurrence  $\mathcal{A}_n(\deg(P_2))$  :

$$\exists Q_2 \in A[X_1, \dots, X_n], P_2(X_1, \dots, X_n) = Q_2(\Sigma_1, \dots, \Sigma_n)$$

On pose  $Q_3(X_1, \dots, X_n) = X_n Q_2(X_1, \dots, X_n)$ , on remarque que  $P_1(X_1, \dots, X_n) = Q_3(\Sigma_1, \dots, \Sigma_n)$ , et on pose également  $\tilde{Q}_1(X_1, \dots, X_n) = Q_1(X_1, \dots, X_{n-1})$ .

D'après la relation (2.1) :

$$\begin{aligned}P(X_1, \dots, X_n) &= Q_3(\Sigma_1, \dots, \Sigma_n) + \tilde{Q}_1(\Sigma_1, \dots, \Sigma_n) \\ &= (Q_3 + \tilde{Q}_1)(\Sigma_1, \dots, \Sigma_n) \\ &= Q_4(\Sigma_1, \dots, \Sigma_n)\end{aligned}$$

Avec  $Q_4 = Q_3 + \tilde{Q}_1$ , on a bien  $Q_4 \in A[X_1, \dots, X_n]$ .

Par récurrence forte, la proposition  $\mathcal{A}_n(p)$  est donc vérifiée pour tout  $p \in \mathbb{N}^*$ .

Par récurrence,  $\mathcal{A}_n$  est vérifiée pour tout  $n \in \mathbb{N}$ .

### Unicité du polynôme $Q$

Soit  $P \in A[X_1, \dots, X_n]$  symétrique.

Supposons qu'il existe deux polynômes  $Q_1$  et  $Q_2$  dans  $A[X_1, \dots, X_n]$  tel que :

$$P(X_1, \dots, X_n) = Q_1(\Sigma_1, \dots, \Sigma_n) = Q_2(\Sigma_1, \dots, \Sigma_n) \quad (2.2)$$

Soit  $\mathcal{H}_n$  la proposition : " $Q(\Sigma_1, \dots, \Sigma_n) = 0 \Rightarrow Q = 0$  dans  $A[X_1, \dots, X_n]$ ".

Pour  $n = 1$  :

Comme vu précédemment, dans  $A[X_1]$ ,  $\Sigma_1 = X_1$ . Donc  $Q(\Sigma_1) = 0 = Q(X_1)$ .

$\mathcal{H}_1$  est vérifiée.

On suppose établie la proposition pour un certain rang  $n - 1 \in \mathbb{N}^*$  :

Soit  $Q \in A[X_1, \dots, X_n]$  tel que  $Q(\Sigma_1, \dots, \Sigma_n) = 0$ . Raisonnons de nouveau par récurrence sur le degré de  $Q$  : on note  $\mathcal{H}_n(p)$  la proposition  $\mathcal{H}_n$  où  $\deg(Q) = p$ .

— Pour  $p = 0$  :

Pour  $Q$  un polynôme constant on a  $Q(\Sigma_1, \dots, \Sigma_n) = c = Q(X_1, \dots, X_n)$ .

Si  $Q(\Sigma_1, \dots, \Sigma_n) = 0$ , alors  $Q = 0$  ;  $\mathcal{H}_n(1)$  est vérifiée.

— On suppose la proposition établie pour un certain  $p \in \mathbb{N}$ .

Notons  $Q(X_1, \dots, X_n) = \sum_{i \in \mathbb{N}^n} a_i X_1^{i_1} \dots X_n^{i_n}$ , avec  $(a_i)_{i \in \mathbb{N}^n} \in A^{\mathbb{N}^n}$  nulle à partir d'un certain rang, et  $i = (i_1, \dots, i_n) \in \mathbb{N}^n$ .

$$\begin{aligned} Q(X_1, \dots, X_n) &= \sum_{i \in \mathbb{N}^n} a_i X_1^{i_1} \dots X_n^{i_n} \\ &= \sum_{\substack{i \in \mathbb{N}^n \\ i_n \neq 0}} a_i X_1^{i_1} \dots X_n^{i_n} + \sum_{\substack{i \in \mathbb{N}^n \\ i_n = 0}} a_i X_1^{i_1} \dots X_n^{i_n} \\ &= X_n \underbrace{\sum_{\substack{i \in \mathbb{N}^n \\ i_n \geq 1}} a_i X_1^{i_1} \dots X_n^{i_n-1}}_{=S(X_1, \dots, X_n)} + \underbrace{\sum_{i \in \mathbb{N}^n} a_i X_1^{i_1} \dots X_{n-1}^{i_{n-1}}}_{=R(X_1, \dots, X_{n-1})} \end{aligned}$$

On obtient donc :

$$\begin{aligned} Q(X_1, \dots, X_n) &= X_n S(X_1, \dots, X_n) + R(X_1, \dots, X_{n-1}) \\ Q(\Sigma_1, \dots, \Sigma_n) &= \Sigma_n S(\Sigma_1, \dots, \Sigma_n) + R(\Sigma_1, \dots, \Sigma_{n-1}) = 0 \end{aligned}$$

En évaluant en  $X_n = 0$ , on a  $(\Sigma_n)_0 = 0$ , d'où :

$$\begin{aligned} Q((\Sigma_1)_0, \dots, (\Sigma_{n-1})_0, 0) &= R((\Sigma_1)_0, \dots, (\Sigma_{n-1})_0) \\ &= R(\Sigma'_1, \dots, \Sigma'_{n-1}) = 0 \end{aligned}$$

De fait,  $R \in A[X_1, \dots, X_{n-1}]$  est tel que  $R(\Sigma'_1, \dots, \Sigma'_n) = 0$ . Donc d'après  $\mathcal{H}_{n-1}$ ,  $R = 0$ . Dès lors,  $Q(X_1, \dots, X_n) = X_n S(X_1, \dots, X_n)$ .

On remarque de plus que  $\deg(S) = \deg(Q) - 1 = p - 1$  et que  $S(\Sigma_1, \dots, \Sigma_n)\Sigma_n = 0$ . Comme  $A[X_1, \dots, X_n]$  est intègre et que  $\Sigma_n \neq 0$ , on a finalement  $S(\Sigma_1, \dots, \Sigma_n) = 0$ ; et par  $\mathcal{H}_n(p - 1)$ ,  $S = 0$ .

Donc  $Q = 0$ .  $\mathcal{H}_n(p)$  est donc vérifiée pour tout  $p \in \mathbb{N}$ . Par principe de récurrence,  $\mathcal{H}_n$  est donc également vérifiée pour tout  $n \in \mathbb{N}^*$ .

Finalemt, notre proposition  $\mathcal{H}_n$  est vérifiée pour tout  $n \in \mathbb{N}^*$ .

Donc, d'après la relation (2.2),  $(Q_1 - Q_2)(\Sigma_1, \dots, \Sigma_n) = 0$ , i.e.  $Q_1 = Q_2$ .

### 2.1.2. Quelques inégalités

Définition 2.1.3 -

Soit  $f$  un polynôme à coefficients complexes.

On pose pour  $t \in \mathbb{C}$ ,  $I_f(t) = \int_{[0;t]} e^{t-u} f(u) du$  et on note  $\tilde{f}$  le polynôme obtenu en remplaçant les coefficients de  $f$  par leur module.

Lemme 2.1.1 -

Soit  $f$  un polynôme à coefficients complexes de degré  $d \in \mathbb{N}$ . On a pour  $t \in \mathbb{C}$  :

$$I_f(t) = e^t \sum_{j=0}^d f^{(j)}(0) - \sum_{j=0}^d f^{(j)}(t) \quad \text{et} \quad |I_f(t)| \leq |t| e^{|t|} \tilde{f}(|t|)$$

*Démonstration.* On fixe  $t \in \mathbb{C}$ .

Montrons la première égalité par récurrence sur  $k \geq 1$  :

$$H_k : I_f(t) = e^t \sum_{j=0}^{k-1} f^{(j)}(0) - \sum_{j=0}^{k-1} f^{(j)}(t) + I_{f^{(k)}}(t)$$

On choisit la paramétrisation  $x \in [0; 1] \mapsto xt$ .

Pour  $k = 1$ , par intégration par parties, on a :

$$\begin{aligned} I_f(t) &= \int_0^1 e^{t-xt} f(xt) t dx \\ &= [-f(xt) e^{t(1-x)}]_0^1 + \int_0^1 f'(xt) e^{t-xt} xt dx \\ &= e^t f(0) - f(t) + I_{f'}(t) \end{aligned}$$

donc  $H_1$  est vraie.

On suppose établie  $H_k$ , pour un certain  $k \geq 1$ . On a par une intégration par parties :

$$\begin{aligned} I_{f^{(k)}}(t) &= \int_0^1 e^{t-xt} f^{(k)}(xt) t dx \\ &= e^t f^{(k)}(0) - f^{(k)}(t) + \int_0^1 e^{t-xt} f^{(k+1)}(xt) dx \\ &= e^t f^{(k)}(0) - f^{(k)}(t) + I_{f^{(k+1)}}(t) \end{aligned}$$

Donc, d'après l'hypothèse de récurrence, on a :

$$I_f(t) = e^t \sum_{j=0}^{k-1} f^{(j)}(0) - \sum_{j=0}^{k-1} f^{(j)}(t) + e^t f^{(k)}(0) - f^{(k)}(t) + I_{f^{(k+1)}}(t)$$

Donc :

$$I_f(t) = e^t \sum_{j=0}^k f^{(j)}(0) - \sum_{j=0}^k f^{(j)}(t) + I_{f^{(k+1)}}(t)$$

Ainsi  $H_{k+1}$  est vraie.

Par le principe de récurrence on a :

$$\forall k \in \mathbb{N}^*, I_f(t) = e^t \sum_{j=0}^{k-1} f^{(j)}(0) - \sum_{j=0}^{k-1} f^{(j)}(t) + I_{f^{(k)}}(t)$$

Dès lors pour  $k = d + 1$  on a  $f^{d+1} = 0$  et on a le résultat souhaité :

$$I_f(t) = e^t \sum_{j=0}^d f^{(j)}(0) - \sum_{j=0}^d f^{(j)}(t)$$

Établissons maintenant l'inégalité. On choisit comme paramétrisation  $x \in [0; 1] \mapsto xt$ . On a alors par l'inégalité triangulaire de l'intégrale :

$$|I_f(t)| = \left| \int_0^1 e^{t-xt} f(xt) t dx \right| \leq \int_0^1 |e^{(1-x)t}| |f(xt)| |t| dx$$

Par l'inégalité triangulaire, sur la somme cette fois, on a pour  $x \in [0; 1]$  si l'on note  $a_k$  les coefficients (complexes) de  $f$  :

$$\begin{aligned} |f(xt)| &= \left| \sum_{k=0}^d a_k (xt)^k \right| \leq \sum_{k=0}^d |a_k| |xt|^k \leq \sum_{k=0}^d |a_k| |t|^k \\ |f(xt)| &\leq \tilde{f}(|t|) \end{aligned}$$

Dès lors, en remarquant que  $\forall z \in \mathbb{C}, |e^z| \leq e^{|z|}$ , on obtient la majoration suivante :

$$\int_0^1 |e^{(1-x)t}| |f(xt)| |t| dx \leq \int_0^1 e^{(1-x)|t|} \tilde{f}(|t|) |t| dx$$

Ainsi par croissance de l'exponentielle, on a :

$$\int_0^1 e^{(1-x)|t|} \tilde{f}(|t|) |t| dx \leq \int_0^1 e^{|t|} \tilde{f}(|t|) |t| dx$$

Ce qui permet de conclure :

$$|I_f(t)| \leq |t| e^{|t|} \tilde{f}(|t|)$$

□

Lemme 2.1.1 -

Soit  $f$  un polynôme à coefficients complexes de coefficient dominant  $l$ , de degré  $d \in \mathbb{N}$ , dont les racines sont majorées en module par  $M$ . On a alors :

$$\forall x \in \mathcal{B}(0, M), \tilde{f}(|x|) \leq |l| \times (2M)^d$$

*Démonstration.* On pose pour  $g \in \llbracket 0; d \rrbracket$ ,  $\sigma_g = \sum_{1 \leq i_1 < \dots < i_g \leq d} x_{i_1} \cdots x_{i_g}$  le  $g$ -ème polynôme symétrique de  $f$  évalué en les racines  $(x_i)_{i \in \llbracket 1; d \rrbracket}$  de  $f$ .

On a alors, en notant  $(a_g)_{g \in \llbracket 0; d \rrbracket}$  les coefficients de  $f$ , par les relations coefficients racines :

$$\sigma_g = (-1)^g \frac{a_{d-g}}{l}$$

Dès lors on a :

$$|a_g| = |l| \times |\sigma_{d-g}|$$

On a par l'inégalité triangulaire et en majorant le module des racines de  $f$  par  $M$  :

$$|\sigma_{d-g}| \leq \sum_{1 \leq i_1 < \dots < i_{d-g} \leq d} |x_{i_1} \cdots x_{i_{d-g}}| \leq \sum_{1 \leq i_1 < \dots < i_{d-g} \leq d} M^{d-g}$$

De plus le cardinal de  $\{(i_1, \dots, i_{d-g}) \mid 1 \leq i_1 < \dots < i_{d-g} \leq d\}$  est  $\binom{d}{d-g}$  car pour déterminer un tel uplet, il faut et il suffit de choisir  $d-g$  entiers de  $\llbracket 1; d \rrbracket$  étant donné que les  $i_k$  sont ordonnés.

Dès lors :

$$|a_g| \leq |l| \times \binom{d}{d-g} M^{d-g}$$

Il vient alors pour  $x \in \mathcal{B}(0, M)$  :

$$\tilde{f}(|x|) = \sum_{g=0}^d |a_g| |x|^g \leq \sum_{g=0}^d |l| \times \binom{d}{d-g} M^{d-g} M^g = |l| \times M^d \sum_{g=0}^d \binom{d}{d-g} = |l| \times M^d \sum_{g=0}^d \binom{d}{g} = |l| \times M^d \times 2^d = |l| \times (2M)^d$$

On conclut alors par le binôme de Newton :

$$\tilde{f}(|x|) \leq |l| \times (2M)^d$$

□

---

## 2.2. Transcendance de $e$

---

### 2.2.1. Résultat

Nous allons à présent montrer le premier résultat souhaité : la transcendance de  $e$ . Pour l'énoncer et le démontrer, nous nous appuierons sur un extrait du livre de Alan BAKER [1].

Théorème 2.2.1 -

|  $e$  est transcendant.

Pour démontrer ce théorème nous raisonnons par l'absurde. Si  $e$  est algébrique, il existe un entier  $n > 0$  et  $q_0, \dots, q_n$  des entiers non tous nuls tels que  $\sum_{k=0}^n q_k e^k = 0$ . On peut supposer  $q_0 \neq 0$  quitte à diviser par une bonne puissance de  $e$ .

Pour la suite de la démonstration, on prend  $p$  un nombre premier, grand, tel que  $p > n$  et  $p > |q_0|$ . On introduit également le polynôme  $f = X^{p-1} (X-1)^p \dots (X-n)^p$ , de degré  $m = (n+1)p - 1$  et  $J = \sum_{k=0}^n q_k I_f(k)$ .

La preuve se déroule comme suit : nous allons dans un premier temps établir une minoration de l'entier  $|J|$ , qui découlera de sa divisibilité par  $(p-1)!$ , et ensuite nous nous appuierons sur les différents lemmes pour établir une majoration de cet entier. Nous obtiendrons alors un encadrement asymptotiquement faux pour  $p$  grand, ce qui conclura la démonstration.

2.2.2. Calcul de  $J$ 

Commençons donc par établir une expression intéressante de  $J$ . Par le Lemme 3, on a :

$$J = \sum_{k=0}^n \sum_{j=0}^m \left( q_k e^k f^{(j)}(0) - q_k f^{(j)}(k) \right)$$

Dès lors en intervertissant les deux sommes finies on obtient :

$$J = \sum_{j=0}^m \sum_{k=0}^n \left( q_k e^k f^{(j)}(0) - q_k f^{(j)}(k) \right)$$

Or par hypothèse  $\sum_{k=0}^n q_k e^k = 0$ , il vient donc :

$$J = - \sum_{j=0}^m \sum_{k=0}^n q_k f^{(j)}(k) = - \left( \sum_{j=0}^m \sum_{k=1}^n q_k f^{(j)}(k) + q_0 f^{(j)}(0) \right)$$

Pour  $j < p$  et  $k \neq 0$  on a  $f^{(j)}(k) = 0$  car  $k$  est une racine de  $f$  de multiplicité  $p$ . De même pour  $j < p-1$ , on a  $f^{(j)}(0) = 0$  car 0 est de multiplicité  $p-1$ .

On a alors :

$$J = - \left( \sum_{j=p}^m \sum_{k=1}^n q_k f^{(j)}(k) + q_0 f^{(j)}(0) \right) - q_0 f^{(p-1)}(0)$$

On pose  $g = X^{p-1}$  et  $h = (X-1)^p \dots (X-n)^p$  de sorte que  $f = gh$ . On a par la formule de Leibniz :

$$f^{(p-1)} = \sum_{k=0}^{p-1} \binom{p-1}{k} g^{(k)} h^{(p-1-k)}$$

et en particulier :

$$f^{(p-1)}(0) = \sum_{k=0}^{p-1} \binom{p-1}{k} g^{(k)}(0) h^{(p-1-k)}(0)$$



0 étant de multiplicité  $p - 1$  pour  $g$ , on a  $\forall k \in \llbracket 0; p - 2 \rrbracket, g^{(k)}(0) = 0$ . Dès lors :

$$f^{(p-1)}(0) = g^{(p-1)}(0) h(0) = (p-1)! (0-1)^p \dots (0-n)^p$$

D'où :

$$f^{(p-1)}(0) = (-1)^{np} (p-1)! (n!)^p$$

Ainsi on a :

$$J = - \left( \sum_{j=p}^m \sum_{k=1}^n q_k f^{(j)}(k) + q_0 f^{(j)}(0) \right) - q_0 (-1)^{np} (p-1)! (n!)^p$$

Cette expression de  $J$  va nous permettre dans la suite de justifier sa non nullité, qui est centrale dans la preuve.

### 2.2.3. Encadrement de $|J|$

Dans l'optique d'établir un encadrement de  $|J|$ , justifions tout d'abord que  $J$  est un entier non nul. Pour ce faire, nous montrerons que  $J$  est un entier divisible par  $(p-1)!$  mais pas par  $p!$ , ce qui découlera du choix de  $p$  premier tel que  $p > |q_0|$  et  $p > n$ .

En notant  $a_l$  les coefficients de  $f$ , on a pour  $j \in \llbracket p; m \rrbracket$  dérivations formelles :

$$\begin{aligned} f^{(j)} &= \sum_{l=0}^{m-j} (l+j)(l+j-1) \dots (l+1) a_{l+j} X^l \\ f^{(j)} &= j! \sum_{l=0}^{m-j} \binom{l+j}{j} a_{l+j} X^l \end{aligned}$$

Les racines de  $f$  sont les éléments de  $\llbracket 0; n \rrbracket$ , et sont donc entières, les  $a_l$  sont donc aussi des entiers par les relations coefficients racines. Dès lors, pour  $j \in \llbracket p; m \rrbracket$  et  $k \in \llbracket 0; n \rrbracket$ ,  $f^{(j)}(k)$  est un entier divisible par  $p!$ .

De plus on a choisi  $p$  premier tel que  $p > n$ , donc  $n!$  n'est pas divisible par  $p$ , de sorte que  $f^{(p-1)}(0)$  est un entier divisible par  $(p-1)!$  mais pas par  $p!$ . Dès lors, vu que  $p > |q_0|$ ,  $p$  ne divise pas  $q_0$  et donc  $p!$  ne divise pas  $q_0 f^{(p-1)}(0)$ .

Ainsi  $J$  est un entier divisible par  $(p-1)!$  mais pas par  $p!$ . On obtient ainsi la non nullité de l'entier  $J$ . En effet s'il était nul,  $p!$  le diviserait. Finalement,  $J$  étant non nul et divisible par  $(p-1)!$ , on a :

$$|J| \geq (p-1)!$$

Il nous reste ainsi à déterminer une majoration de  $|J|$ . Pour ce faire, nous allons nous appuyer sur les différents lemmes montrés précédemment, et plus particulièrement sur le Lemme 3 et le Lemme 4.

On rappelle que par définition :

$$J = \sum_{k=0}^n q_k I_f(k)$$

On a alors, par l'inégalité triangulaire puis le Lemme 3 :

$$|J| \leq \sum_{k=0}^n |q_k| |I_f(k)| \leq \sum_{k=0}^n |q_k| k e^k \tilde{f}(k)$$

Dès lors, vu que  $f$  est unitaire et que ses racines sont les éléments de  $\llbracket 0; n \rrbracket$ , et donc majorées en valeur absolue par  $n$ , on a par le Lemme 4 :

$$|J| \leq \sum_{k=0}^n |q_k| k e^k (2n)^{(n+1)p-1} = \sum_{k=0}^n |q_k| k e^k \frac{(2n)^{(n+1)p}}{2n} \leq \frac{1}{2} \left( \sum_{k=0}^n |q_k| e^k \right) c^p$$

avec  $c = (2n)^{n+1}$ .

Finalement on a montré que pour tout  $p$  premier tel que  $p > |q_0|$  et  $p > n$  on a l'encadrement suivant :

$$(p-1)! \leq |J| \leq \frac{1}{2} \left( \sum_{k=0}^n |q_k| e^k \right) c^p$$

Ceci est absurde asymptotiquement pour  $p$  assez grand, car  $c^p = o((p-1)!)$

Ainsi  $e$  est transcendant.

---

## 2.3. Transcendance de $\pi$

---

### 2.3.1. Résultat

Nous allons donc maintenant montrer le deuxième résultat souhaité : la transcendance de  $\pi$ . Nous nous appuyerons sur la preuve présentée dans le livre d'Alan BAKER [1].

Théorème 2.3.1 -

|  $\pi$  est transcendant.

Nous allons procéder de la même manière que pour  $e$ , par l'absurde. Supposons que  $\pi$  soit algébrique. Commençons par montrer que  $i\pi$  est alors, lui aussi, algébrique.

Lemme 2.3.1 -

| Si  $\alpha \in \mathbb{C}$  est algébrique,  $i\alpha$  l'est aussi.

*Démonstration.* Si  $\alpha$  est algébrique, il existe  $P \in \mathbb{Z}[X]$ ,  $P \neq 0$  tel que  $P(\alpha) = 0$ .

Notons  $P(X) = a_0 + a_1X + \dots + a_nX^n$ , avec  $(a_k)_{k \in \llbracket 0; n \rrbracket} \in \mathbb{Z}^{n+1}$ .

Soit  $Q(X) = P(iX)P(-iX) \in \mathbb{C}[X]$  ; on a  $Q \neq 0$ .

On a que :

$$\begin{aligned}
 Q(X) &= \left( \sum_{k=0}^n a_k i^k X^k \right) \left( \sum_{k=0}^n a_k (-i)^k X^k \right) \\
 &= \sum_{m=0}^{2n} \left( \sum_{k=0}^m a_k a_{m-k} i^k (-i)^{m-k} \right) X^m \quad (\text{en posant pour } k > n \ a_k = 0) \\
 &= \sum_{m=0}^{2n} \underbrace{\left( i^m \left( \sum_{k=0}^m a_k a_{m-k} (-1)^{m-k} \right) \right)}_{=b_m} X^m
 \end{aligned}$$

De plus, pour  $m \in \llbracket 0; 2n \rrbracket$  :

1<sup>er</sup> cas : Si  $m \equiv 0[2]$

$$b_m = \underbrace{i^m}_{\pm 1} \underbrace{\sum_{k=0}^m a_k a_{m-k} (-1)^{m-k}}_{\in \mathbb{Z}} \in \mathbb{Z}$$

2<sup>eme</sup> cas : Si  $m \equiv 1[2]$  ;  $m = 2p + 1$ ,  $p \in \mathbb{N}$  :

$$\begin{aligned}
 b_m &= i^m \left( \sum_{k=0}^{2p+1} a_k a_{m-k} (-1)^{m-k} \right) \\
 &= i^m \left( \sum_{k=0}^p a_k a_{m-k} (-1)^{m-k} + \sum_{k=p+1}^{2p+1} a_k a_{m-k} (-1)^{m-k} \right) \\
 &= i^m \left( \sum_{k=0}^p a_k a_{m-k} (-1)^{m-k} + \sum_{j=0}^p a_{m-j} a_j (-1)^j \right) \\
 &= i^m \left( \sum_{k=0}^p a_k a_{m-k} \underbrace{\left( (-1)^{m-k} + (-1)^k \right)}_{=0} \right) = 0
 \end{aligned}$$

Donc  $Q \in \mathbb{Z}[X]$  et  $Q(i\alpha) = P(-\alpha)P(\alpha) = 0$ . Donc  $i\alpha$  est algébrique.

□

Nous avons donc, par hypothèse, que  $i\pi$  est algébrique. Notons  $P$  son polynôme minimal. Soient  $d \in \mathbb{N}^*$  le degré de  $P$ ,  $l \in \mathbb{Z}$  son coefficient dominant et  $\theta_1 = i\pi, \theta_2, \dots, \theta_d$  ses racines :

$$P = l \prod_{i=1}^d (X - \theta_i) \in \mathbb{Z}[X]$$

### 2.3.2. Autour de l'égalité d'Euler

Comme  $\theta_1 = i\pi$ , on a l'égalité d'Euler  $e^{\theta_1} = -1$ , ainsi :

$$\prod_{i=1}^d (1 + e^{\theta_i}) = 0$$

c'est-à-dire, en développant le produit :

$$\sum_{k=0}^d \sum_{1 \leq i_1 < \dots < i_k \leq d} e^{\theta_{i_1} + \dots + \theta_{i_k}} = 0$$

Ce produit est, finalement, une somme de  $e^{\vartheta}$  où  $\vartheta \in \Omega$ , avec  $\Omega$  le multi-ensemble contenant toutes les sommes de  $(\theta_i)_{1 \leq i \leq n}$  possibles, apparaissant chacun au plus une fois.

De plus, on constate que parmi ces exponentielles, il en existe au moins une égale à 1 (pour  $\vartheta = 0$ ) et au moins une différente de 1 (pour  $\vartheta = \theta_1$ ,  $e^{\theta_1} = -1$ ). Il y a donc exactement  $\alpha_1, \dots, \alpha_n \in \Omega$ , avec  $n \in \mathbb{N}^*$ , représentant, comptés avec multiplicité, les  $\vartheta$  d'exponentielle différentes de 1. Finalement, on a  $q \in \mathbb{N}^*$  tel que :

$$q + e^{\alpha_1} + \dots + e^{\alpha_n} = 0 \quad (2.3)$$

où  $q = \sum_{k=0}^d \binom{d}{k} - n = 2^d - n$ , correspond au nombre de  $\vartheta$  d'exponentielle égale à 1.

### 2.3.3. Calcul de $J$

Considérons maintenant  $p$  un nombre premier grand, et le polynôme  $f = l^{np} X^{p-1} (X - \alpha_1)^p \dots (X - \alpha_n)^p$ , de degré  $m = (n+1)p - 1$ .

Posons  $I_f(t) = \int_{[0;t]} e^{t-u} f(u) du$ , pour  $t \in \mathbb{C}$ , et  $J = \sum_{i=1}^n I_f(\alpha_i)$ . Nous allons, par la suite, démontrer que  $J$  est non nul, et trouver un encadrement asymptotiquement faux pour  $p$  premier assez grand.

Notons  $\gamma : x \in [0; 1] \mapsto xt$  et

$$I_f(t) = \int_{\gamma} e^{t-u} f(u) du = \int_0^1 e^{t-xt} f(xt) t dx$$

Par le Lemme 3 :

$$I_f(t) = e^t \sum_{j=0}^m f^{(j)}(0) - \sum_{j=0}^m f^{(j)}(t) \quad (2.4)$$

Soit  $J = \sum_{k=1}^n I_f(\alpha_k)$ .

On a avec l'égalité (2.4) :

$$\begin{aligned} J &= \sum_{k=1}^n \left( e^{\alpha_k} \sum_{j=0}^m f^{(j)}(0) - \sum_{j=0}^m f^{(j)}(\alpha_k) \right) \\ &= \sum_{j=0}^m \sum_{k=1}^n e^{\alpha_k} f^{(j)}(0) - \sum_{j=0}^m \sum_{k=1}^n f^{(j)}(\alpha_k) \end{aligned}$$

Et, d'après l'égalité (2.3) :

$$J = -q \sum_{j=0}^m f^{(j)}(0) - \sum_{j=0}^m \sum_{k=1}^n f^{(j)}(\alpha_k)$$

2.3.4. Étude de  $J$ 

$$\text{Étude de } \sum_{j=0}^m \sum_{k=1}^n f^{(j)}(\alpha_k)$$

Nous allons d'abord nous intéresser au terme de droite dans l'expression de  $J$ . Montrons que les  $P_j = \sum_{k=1}^n f^{(j)}(\alpha_k)$  sont des polynômes symétriques en  $l\alpha_1, \dots, l\alpha_n$ .

Pour  $j < p$ ,  $f^{(j)}(\alpha_k) = 0$  pour tout  $k$ , ainsi  $P_j = 0$  pour  $j < p$ . De plus, si  $j \geq p$ , on a dérivé  $j$  fois  $f$ ; il va donc rester, quand on calcule  $f^{(j)}$ , une somme de produits de  $m - j = (n+1)p - 1 - j$  polynômes de degrés 1, avec  $(n+1)p - 1 - j < np$ . Or  $l^{np}$  est en facteur de ces polynômes, on peut donc multiplier chacun des polynômes, de la forme  $(X - \alpha_i)$  ou  $X$ , par un  $l$ . Chaque  $(X - \alpha_i)$  devient un  $(lX - l\alpha_i)$  et chaque  $X$  devient un  $lX$ . Ainsi  $f^{(j)}(\alpha_k)$  peut être vu comme un polynôme à coefficients entiers en  $l\alpha_1, \dots, l\alpha_n$ .

Finalement, pour tout  $j$ ,  $\sum_{k=1}^n f^{(j)}(\alpha_k) = P_j \in \underline{\mathbb{Z}[l\alpha_1, \dots, l\alpha_n]}$

Si on prend  $\sigma$  une permutation de  $\{1, \dots, n\}$  :

$$P_j(l\alpha_{\sigma(1)}, \dots, l\alpha_{\sigma(n)}) = \sum_{k=1}^n f^{(j)}(\alpha_{\sigma(k)}) = \sum_{k'=1}^n f^{(j)}(\alpha_{k'})$$

Donc  $P_j$  est symétrique.

On peut donc appliquer le théorème 1 fondamental des polynômes symétriques à chaque  $P_j \in \mathbb{Z}[l\alpha_1, \dots, l\alpha_n]$ . Il existe un polynôme  $Q_j$  à coefficients entiers, à  $n$  indéterminées, tel que :

$$P_j(l\alpha_1, \dots, l\alpha_n) = Q_j(\Sigma_1, \dots, \Sigma_n)$$

où les  $\Sigma_1, \dots, \Sigma_n$  sont les polynômes symétriques élémentaires en  $l\alpha_1, \dots, l\alpha_n$ .

Nous allons montrer que les polynômes symétriques élémentaires  $\Sigma_1, \dots, \Sigma_n$  sont des entiers.

Première étape :  $\theta_1, \dots, \theta_d$  sont les racines du polynôme  $P$ , le polynôme minimal de  $i\pi$ , tel que :

$$P = l \prod_{i=1}^d (X - \theta_i) \in \mathbb{Z}[X]$$

Si on note  $\sigma_1, \dots, \sigma_d$  les polynômes symétriques élémentaires en  $\theta_1, \dots, \theta_d$ . On remarque que :

$$P = lX^d + \sum_{i=1}^d (-1)^i l\sigma_i X^{d-i} \in \mathbb{Z}[X]$$

Donc  $l\sigma_1, \dots, l\sigma_d$  sont entiers.

Deuxième étape : On considère  $\Omega = \{\alpha_1, \dots, \alpha_n, \alpha_{n+1}, \dots, \alpha_{2d}\}$ , le multi-ensemble, défini dans la partie 2.3.2, contenant toutes les sommes possibles de  $\theta_1, \dots, \theta_n$ , apparaissant chacun au plus une fois. Dans  $\Omega$ , on rappelle que  $\alpha_1, \dots, \alpha_n$  sont les éléments non nuls. Considérons le polynôme :

$$S = (X - l\alpha_1) \dots (X - l\alpha_{2d})$$

En développant  $S$ , on obtient comme coefficients, au signe près, les polynômes symétriques élémentaires en  $l\alpha_1, \dots, l\alpha_{2d}$ , notés  $A_i$ .

Chaque  $A_i$  peut également être vu comme un polynôme à coefficients entiers en les  $l\theta_1, \dots, l\theta_d$ . Nous allons voir qu'il est symétrique en ces variables. Si on prend une permutation de  $\{1, \dots, n\}$ , étant donné que l'ensemble  $\Omega$  contient toutes les sommes possibles de  $\theta_1, \dots, \theta_d$  apparaissant chacun au plus une fois, permuter les  $l\theta_1, \dots, l\theta_d$  dans  $A_i$  revient à permuter les  $l\alpha_1, \dots, l\alpha_{2d}$ . Finalement, comme  $A_i$  est symétrique en  $l\alpha_1, \dots, l\alpha_{2d}$ , on retrouve que  $A_i$  l'est également en les  $l\theta_1, \dots, l\theta_d$ .

Finalement, les  $A_i$  sont des polynômes symétriques en  $l\theta_1, \dots, l\theta_d$  à coefficients entiers, d'après le théorème 1 fondamental sur les polynômes symétriques, il existe

$B_i \in \mathbb{Z}[X_1, \dots, X_d]$  tels que :

$$A_i(l\theta_1, \dots, l\theta_d) = B_i(l\sigma_1, \dots, l\sigma_d)$$

On a montré précédemment que  $l\sigma_1, \dots, l\sigma_d \in \mathbb{Z}$ , donc les  $A_i \in \mathbb{Z}$ , et donc  $S \in \mathbb{Z}[X]$ .

Troisième étape :  $S = (X - l\alpha_1) \dots (X - l\alpha_{2d}) = (X - l\alpha_1) \dots (X - l\alpha_n) X^{2d-n} \in \mathbb{Z}[X]$

En développant le polynôme  $S$ , les  $\Sigma_1, \dots, \Sigma_n$ , correspondant aux polynômes symétriques élémentaires en les  $l\alpha_1, \dots, l\alpha_n$ , sont au signe près les coefficients de  $S$ , ils sont donc entiers.

En conclusion, les  $P_j = \sum_{k=1}^n f^{(j)}(\alpha_k)$  sont des entiers.

Nous allons maintenant montrer que  $p!$  divise  $\sum_{j=0}^m \sum_{k=1}^n f^{(j)}(\alpha_k) = \sum_{j=0}^m P_j$ .

Comme  $f = X^{P-1}((lX - l\alpha_1) \dots (lX - l\alpha_n))^P$ , en notant  $a_i$  ces coefficients, on obtient que chaque  $a_i$  est, au signe près, le produit de polynômes symétriques élémentaires en les  $l\alpha_1, \dots, l\alpha_n$ , multiplié par une puissance de  $l$ . Nous avons montré précédemment, que ces polynômes étaient entiers. Ces  $a_i$  sont donc entiers, et on a pour  $j \in \llbracket p; m \rrbracket$  dérivations formelles :

$$\begin{aligned} f^{(j)} &= \sum_{i=0}^{m-j} (i+j)(i+j-1) \dots (i+1) a_{i+j} X^i \\ f^{(j)} &= j! \sum_{i=0}^{m-j} \binom{i+j}{j} a_{i+j} X^i \end{aligned}$$

Ainsi, les coefficients de  $f^{(j)}$  sont divisibles par  $j!$  et donc par  $p!$  pour  $j \geq p$ . Comme les  $\alpha_k$  sont racines d'ordre  $p$  de  $f$ , on a :

$$\sum_{j=0}^m \sum_{k=1}^n f^{(j)}(\alpha_k) = \sum_{j=p}^m \sum_{k=1}^n f^{(j)}(\alpha_k)$$

Pour résumer, d'après ce que l'on vient d'énoncer,  $\sum_{j=p}^m \sum_{k=1}^n f^{(j)}(\alpha_k)$  est un polynôme symétrique à coefficients entiers en les  $l\alpha_1, \dots, l\alpha_n$ , dont chaque coefficient est divisible par  $p!$ . On peut donc écrire  $\sum_{j=p}^m \sum_{k=1}^n f^{(j)}(\alpha_k) = p!T$  avec  $T$  un polynôme symétrique à coefficients entiers en les  $l\alpha_1, \dots, l\alpha_n$ . Par le

théorème fondamental sur les polynômes symétriques 1, il existe  $Q \in \mathbb{Z}[X_1, \dots, X_n]$  tel que :

$$T(l\alpha_1, \dots, l\alpha_n) = Q(\Sigma_1, \dots, \Sigma_n)$$

où  $\Sigma_1, \dots, \Sigma_n$  sont les polynômes symétriques élémentaires en les  $l\alpha_1, \dots, l\alpha_n$ . Or les  $\Sigma_1, \dots, \Sigma_n$  sont des entiers, donc il vient que  $Q(\Sigma_1, \dots, \Sigma_n)$  est un entier.

$$\text{Finalement, } p! \text{ divise } \sum_{j=0}^m \sum_{k=1}^n f^{(j)}(\alpha_k) = \sum_{j=0}^m P_j.$$

$$\text{Étudie } f^{(j)}(0)$$

Nous allons maintenant nous intéresser au terme de gauche dans l'expression de  $J$ . On rappelle que :

$$J = -q \sum_{j=0}^m f^{(j)}(0) - \sum_{j=0}^m \sum_{k=1}^n f^{(j)}(\alpha_k)$$

Posons  $f = hg$  avec  $h = l^{np} (X - \alpha_1)^p \dots (X - \alpha_n)^p$  et  $g = X^{p-1}$ . Nous remarquons que  $h = ((lX - l\alpha_1) \dots (lX - l\alpha_n))^p$ , et nous avons montré précédemment que les polynômes symétriques élémentaires en  $l\alpha_1, \dots, l\alpha_n$  étaient entiers. Donc  $h$  est un polynôme à coefficients entiers. Dès lors,  $f^{(j)}(0) \in \mathbb{Z}$ .

- Pour  $j < p - 1$ ,  $f^{(j)}(0) = 0$

- Pour  $j = p - 1$ , on montre, en utilisant la formule de Leibniz appliquée à  $f = hg$  que :

$$f^{(p-1)}(0) = \binom{p-1}{p-1} l^{np} (p-1)! (-\alpha_1)^p \dots (-\alpha_n)^p = (-1)^p (p-1)! l^{p(n-1)} \underbrace{(l\alpha_1 \dots l\alpha_n)^p}_{=\Sigma_n \in \mathbb{Z}}$$

Donc  $f^{(p-1)}(0)$  est un entier divisible par  $(p-1)!$ . De plus, en prenant  $p$  assez grand, c'est-à-dire  $p > l$ ,  $p > \Sigma_n$  et  $p > q$ ,  $qf^{(p-1)}(0)$  est un entier divisible par  $(p-1)!$  mais pas par  $p!$ .

- Pour  $j > p - 1$ , on a de nouveau, par la formule de Leibniz :

$$f^{(j)}(0) = \sum_{k=p-1}^j \binom{j}{k} g^{(k)}(0) h^{(j-k)}(0) = \binom{j}{p-1} (p-1)! h^{(j-p+1)}(0)$$

On remarque que  $h = ((lX - l\alpha_1) \dots (lX - l\alpha_n))^p = h_1^p$ , avec  $h_1 \in \mathbb{Z}[X]$ . On a  $h' = ph_1' h_1^{(p-1)} = ph_2$ , avec encore  $h_2 \in \mathbb{Z}[X]$ . Comme  $j > p - 1$ , c'est-à-dire  $j - p \geq 0$ , on peut écrire  $h^{(j-p+1)} = (h')^{(j-p)} = ph_2^{(j-p)}$ . Finalement, on a :

$$f^{(p-1)}(0) = \binom{j}{p-1} (p-1)! p h_2^{(j-p)}(0) = \binom{j}{p-1} p! \underbrace{h_2^{(j-p)}(0)}_{\in \mathbb{Z}}$$

Donc  $p!$  divise  $f^{(j)}(0)$ , pour  $j > p - 1$ .

Ainsi, on a montré que dans la somme  $q \sum_{j=0}^m f^{(j)}(0)$ , tout les termes étaient divisibles par  $(p-1)!$

et un seul et unique terme n'était pas divisible par  $p!$ . Dès lors,  $q \sum_{j=0}^m f^{(j)}(0)$  est un entier divisible par  $(p-1)!$  mais pas par  $p!$ .

### 2.3.5. Encadrement de $|J|$

Pour finir,  $J$  est un entier divisible par  $(p-1)!$  et pas par  $p!$ . Il est donc non nul. On a alors :

$$|J| \geq (p-1)!$$

Or, d'après l'inégalité triangulaire et le Lemme 3 :

$$|J| \leq \sum_{k=0}^n |I_f(\alpha_k)| \leq \sum_{k=0}^n |\alpha_k| e^{|\alpha_k|} \tilde{f}(|\alpha_k|)$$

Le coefficient dominant de  $f$  est  $l^{np}$  et on note  $\alpha = \max\{|\alpha_k|, k \in \llbracket 0; n \rrbracket\}$  un majorant des racines de  $f$ . On a alors par le Lemme 4 :

$$|J| \leq \sum_{k=0}^n |\alpha_k| e^{|\alpha_k|} |l|^{np} (2\alpha)^{(n+1)p-1} \leq \sum_{k=0}^n \alpha e^{|\alpha_k|} |l|^{np} (2\alpha)^{(n+1)p-1} \leq \frac{1}{2} \left( \sum_{k=0}^n e^{|\alpha_k|} \right) c^p$$

avec  $c = |l|^n (2\alpha)^{n+1}$ .

Finalement, on a montré que pour tout  $p$  premier assez grand, on a l'encadrement suivant :

$$(p-1)! \leq |J| \leq \frac{1}{2} \left( \sum_{k=0}^n e^{|\alpha_k|} \right) c^p$$

Ceci est absurde asymptotiquement pour  $p$  grand, car  $c^p = o((p-1)!)$

Ainsi  $\pi$  est transcendant.

---

## Liste des références

---

Alan Baker. *Transcendental number theory*, chapter The Origins. Cambridge University Press, 1975.

Théorème de structure des polynômes symétriques, leçon pour l'agrégation, 2020.

X. Gourdon. *Algèbre*, page 89. Ellipses, 2008.



# Rapport de stage de fin de licence



GABAUT Auriane  
Encadré par Pasquale DI FATTA  
Mai à Juin 2021

## Sommaire

|                                                            |           |
|------------------------------------------------------------|-----------|
| <b>Introduction</b>                                        | <b>37</b> |
| <b>3.1 Première semaine du stage : phase d'observation</b> | <b>38</b> |
| 3.1.1 Organisation de la classe                            | 38        |
| 3.1.2 Attendus                                             | 38        |
| <b>3.2 Tâches confiées</b>                                 | <b>39</b> |
| 3.2.1 Interrogations écrites                               | 39        |
| 3.2.2 DM et TD                                             | 40        |
| 3.2.3 Colles                                               | 44        |
| 3.2.4 Correction d'exercices et démonstrations             | 48        |
| <b>Conclusion</b>                                          | <b>55</b> |

## Introduction

Ce stage a été effectué dans le cadre du magistère de mathématiques en L3. J'ai eu l'opportunité de le réaliser au Lycée Georges Clémenceau, à Reims, en classe de MPSI. Il a été encadré par M. Di Fatta, professeur de mathématiques en MPSI2. Ce stage s'est déroulé du 17 mai au 12 juin 2021.

L'objectif de ce stage était, pour moi, de découvrir au mieux le métier d'enseignant au sein d'une classe prépa. Mes attentes pour ce stage étaient de confirmer, ou d'infirmer, mon projet de devenir professeur.

---

### 3.1. Première semaine du stage : phase d'observation

---

Cette première partie correspond majoritairement à l'observation faite de la classe. J'ai d'abord eu l'occasion d'observer le déroulement des cours et de la semaine, non pas en me concentrant sur le contenu du cours mais sur sa forme. Lors de ma venue, les chapitres traités ont été les probabilités, les matrices, les variables aléatoires, les déterminants et le début du chapitre sur les espaces préhilbertiens.

#### 3.1.1. Organisation de la classe

La classe dans laquelle j'ai effectué mon stage est une classe de 47 élèves. L'ambiance y est globalement bonne et propice à l'entraide, ce que leurs professeurs encourage vivement. Les cours de la semaine se divisent en deux parties :

- les cours : 11h par semaine
- les TD : 2h par semaine

Lors des cours, le professeur avance sur le chapitre en alternant explication de cours et rédaction de démonstration. Le cours est donné en amont, sous forme de polycopié à trous. Les élèves doivent compléter les propositions et théorèmes, et recopier intégralement les démonstrations. Ce système permet au professeur de garder l'attention sur lui lors des explications et de ne pas avoir d'élève occupé à recopier, tout en s'assurant qu'ils ne se reposent pas seulement sur le polycopié, mais suivent bel et bien le cours. A chaque début de cours, le professeur corrige quelques exercices donnés à faire à la maison, ou des exemples d'applications du cours.

Pour les TD, les élèves sont répartis en deux groupes, A et B. Dans chaque groupe, les élèves sont divisés en groupe de 4 et ont une liste d'exercices, entre 5 et 10 par semaine, à chercher au cours des 2h. Le premier groupe qui arrive à 20 exercices résolus, fixe les notes ; ils obtiennent 20, le deuxième groupe à avoir résolu le plus d'exercices obtient 19 etc. Ceci permet à chaque élève de chercher réellement les exercices, et de ne pas être passif lors de la correction, mais aussi d'être aidé par des membres plus forts de ce groupe. La note obtenue à la fin permet de s'assurer que chaque étudiant aie une motivation immédiate à travailler.

Pendant la recherche des exercices, le professeur passe dans les groupes 1 à 1, en restant entre 5 et 15 minutes avec chaque groupe, selon le niveau de difficulté. Cela lui permet de contrôler le travail effectué, de répondre aux questions et de valider la résolution d'un exercice.

#### 3.1.2. Ateliers

En plus de ces cours, il est prévu dans l'emploi du temps un devoir sur table de 4h toutes les 3 semaines, type concours, qui permettent de préparer les étudiants pour ces derniers. Ils ont également

à chaque début de période, c'est à dire entre chaque vacances, une interrogation écrite de cours, d'une heure, synthétisant tout ce qui a été vu à la période précédente. Il s'agit d'un contrôle de cours survolant l'intégralité des chapitres précédents. Toutes les propositions importantes sont demandées ainsi que les théorèmes. Cette dernière oblige les étudiants à ne pas laisser le cours de côté pendant les vacances. A cela s'additionne une petite interrogation écrite de 15 minutes chaque début de semaine, avec quelques questions de cours sur la semaine d'avant et un exercice d'application.

A la maison, les étudiants ont 3 exercices de la feuille de TD à faire chaque soir, ils finissent également les exercices d'application du cours rencontrés. Ils doivent également préparer un DM de mathématiques chaque semaine où ils n'ont pas DS de mathématiques.

Les étudiants ont une colle de mathématiques d'une heure toutes les semaines, le programme de colle est donné le vendredi précédent. Ils doivent maîtriser 5 démonstrations choisis à l'avance, seules celles si peuvent être demandées. Ils répondent ensuite à deux questions de cours puis résolvent deux exercices.

---

## 3.2. Tâches confiées

---

Cette deuxième partie synthétise les différentes approches que j'ai pu essayer et leurs mises en application au cours de mon stage. J'ai eu, en effet, la chance d'avoir eu beaucoup de tâches variées à effectuer. J'ai pu corriger des petites interrogations écrites et quelques copies de DM, j'ai pu faire plusieurs colles et encadrer des TD. J'ai également pu corriger des exercices en début de cours ainsi que des démonstrations.

### 3.2.1. Interrogations écrites

L'une des premières tâches qui m'a été confié, puis que j'ai demandé volontairement à renouveler chaque semaine, a été la correction des petites interrogations hebdomadaires. J'ai pu profiter d'une confiance complète de la part de mon encadrant sur le barème et la correction choisie, j'ai eu ses conseils et instructions pour la première interrogation puis une complète autonomie m'a été laissée.

Les petites interrogations hebdomadaires étaient notées sur 2 points, un pour le cours, un pour l'exercice. Au bout de 10 interrogations, la somme des notes donne la note finale. Les interrogations de début de période sont notées sur 20 points

#### Interrogation hebdomadaires

L'encadrant m'a d'abord fourni les interrogations n°23, présenté en annexe Figure 10 lors de la deuxième semaine. J'ai pu alors d'abord établir un barème : j'ai décidé de mettre 0.2 point par réponses demandées sur le cours ; le barème pour l'exercice était de 0.2 point pour la question 1 (0.1 pour la

linéarité et 0.1 pour avoir précisé de  $E$  à valeurs dans  $E$ ), 0.2 point pour la question 2, et 0.6 point pour la dernière question (0.4 pour la bijectivité et 0.2 pour le calcul de l'inverse). Après avoir corrigé une majorité des copies, j'ai décidé d'enlever 0.1 supplémentaire à tout ceux n'ayant pas précisé où doivent être choisis les matrices inversibles dans les formules de changement de bases, j'ai également décidé de ne mettre que 0.1 point à ceux ayant juste inversé la matrice associée à l'endomorphisme pour calculer l'inverse. En effet, l'exercice ayant été traité en cours, il était attendu de donner une expression explicite de  $\varphi$ , par exemple en fonction des coordonnées de  $P = aX^2 + bX + c \in \mathbb{R}^2[X]$ .

Pour les trois autres interrogations, en Figure 9, 11 et 12, la correction de l'exercice ne reposait pas sur une répartition des points sur telle ou telle élément de réponse. J'ai donné la note de 1/1 pour l'exercice lorsque la réponse était correcte et justifiée, à défaut, 0.2 point était retiré à chaque erreur.

### *Interrogation de début de période*

La rentrée des vacances de mai et le début de mon stage coïncidant, j'ai pu corrigé une interrogation écrite de début de période. L'énoncé de l'interrogation est fourni en annexe A. Le barème a été établi avec le professeur, de sorte à vérifier que les points étaient bien répartis, et concordaient avec ses attentes. La correction de ses interrogations étaient les premières que j'ai eu à réaliser, j'ai alors pu me familiariser avec les derniers chapitres étudiés, de sorte à répondre aux questions des étudiants en vu des DS et DM sans déborder sur le programme de deuxième année.

### *3.2.2. DM et TD*

#### *TD encadrés*

La correction des TD est importante pour faire progresser les étudiants. C'est l'occasion d'appliquer le cours, de s'exercer et de vérifier la bonne compréhension. En temps normal, le professeur passe dans les groupes pendant l'heure un à un, en général 1 à 2 fois par TD pour chaque groupe. Ma présence a permis de faire un passage de plus puisque, plutôt que de regarder l'encadrant aider chaque groupe d'étudiants, j'ai pu faire une rotation dans les groupes. J'ai aidé les étudiants lors de 3 TD.

**TD du 10 mai** Les exercices donnés lors de ce TD concernaient les probabilités ainsi que des révisions d'algèbre linéaire :

**Exercice 1.** Si  $E = \mathbb{R}[X]$  et  $A = X^3 + 1 \in E$ , on note  $I_A = \{P \in E; A|P\}$ .

1. Montrer que  $I_A$  est un sous-espace vectoriel de  $E$  et que  $I_A \oplus \mathbb{R}_2[X] = E$
2. Trouver le projeté de  $X^4 + 1$  sur  $\mathbb{R}_2[X]$  parallèlement à  $I_A$ .
3. Trouver le symétrique de  $X^4 + 1$  par rapport à  $I_A$  et parallèlement à  $\mathbb{R}_2[X]$ .

**Exercice 2.** Montrer que  $A = \{y \in C^2(\mathbb{R}); x^2 y'' - xy' + y = x^2\}$  est un sous-espace affine de  $C^2(\mathbb{R})$ .

**Exercice 3.** Si  $E$  est un  $\mathbb{K}$ -espace vectoriel,  $a \in E$  non nul, et  $u$  une forme linéaire non nulle sur  $E$ . On pose si  $x \in E$  :  $f(x) = u(x)a$

1. Montrer que  $f$  est un endomorphisme de  $E$  et que  $Im(f) = Vect(a)$ .
2. Déterminer une C.N.S. pour que  $f$  soit un projecteur de  $E$ .
3. Calculer pour tout  $k \in \mathbb{N}$ ,  $f^k$ .

**Exercice 4.** Soient  $E$  un  $\mathbb{K}$ ev et  $p, q$  deux projecteurs de  $E$  qui commutent.

1. Montrer que  $poq$  est un projecteur de  $E$ .
2. Montrer que  $Ker(poq) = Ker(p) + Ker(q)$  et  $Im(poq) = Im(p) \cap Im(q)$

**Exercice 5.** On a deux dés à six faces, un parfaitement équilibré et un second truqué qui amène 6 avec une probabilité de  $1/2$ , et les autres chiffres de manière équiprobable. On choisit au hasard un de ces deux dés et on le lance.

1. On obtient 6, quelle est la probabilité que le dé choisi soit le dé équilibré ?
2. On obtient 5, même question.

**Exercice 6.** Une urne contient 15 boules : 1 noire, 5 blanches et 9 rouges.

1. On tire simultanément et au hasard 3 boules dans cette urne, quelle est la probabilité des événements suivants :
  - (a) Le tirage est tricolore.
  - (b) Le tirage contient une boule noire et au moins une boule rouge.
  - (c) Le tirage contient trois boules de même couleur.
2. Reprendre les questions avec un tirage successif avec remise.

**TD du 31 mai** Les exercices donnés lors de ce TD concernaient les matrices et les variables aléatoires :

**Exercice 1.** Déterminer le rang de la matrice  $C = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & 1 & -2 \\ 1 & 2 & 4 & -1 & 5 \\ 3 & 9 & 11 & -2 & 19 \end{pmatrix}$

**Exercice 2.** Déterminer le rang de la matrice  $B = \begin{pmatrix} 1 & a & a & a \\ a & 1 & a & a \\ a & a & 1 & a \\ a & a & a & 1 \end{pmatrix}$

**Exercice 3.** Montrer que  $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$  et  $B = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$  sont les matrices d'un même automorphisme de  $\mathbb{R}^2$  relativement à des bases différentes.

**Exercice 4.** Si  $A \in \mathcal{M}_n(\mathbb{K})$ , montrer que  $\text{Tr}({}^tAA) \geq 0$  et que  $\text{Tr}({}^tAA) = 0$  ssi  $A = 0_n$ .

**Exercice 5.** Pour chaque question, reconnaître la loi de  $X$  et préciser ses paramètres :

1. On effectue un tirage dans une urne qui contient  $n$  boules numérotées de 1 à  $n$ , on note  $X$  le nombre obtenu.
2. Une urne contient  $n$  boules vertes,  $p$  boules rouges, on effectue des tirages successifs avec remise de  $k$  boules et on note  $X$  le nombre de boules vertes tirées.
3. On aligne les 32 cartes d'un jeu de cartes faces cachées sur une table, on retourne les cartes de gauche à droite jusqu'à obtenir le valet de trèfle et  $X$  donne le nombre de cartes retournées.
4. On répartit au hasard 100 boules dans 3 urnes :  $X$  compte le nombre de boules dans la première urne.
5. Lors d'une visite médicale de  $n$  étudiants tous de taille différente, le médecin les reçoit un par un de manière aléatoire et les mesure. On note  $X$  le rang dans l'ordre de passage du plus petit des étudiants.
6. Dans un QCM de  $n$  questions avec  $r$  réponses pour chaque question dont une seule est correcte,  $X$  compte le nombre de bonnes réponses.

**Exercice 6.**  $X$  est une variable aléatoire réelle sur  $\Omega$  fini, on appelle fonction de répartition de  $X$  notée  $F_X$  la fonction définie sur  $\mathbb{R}$  par  $F_X(t) = \mathbb{P}(X \leq t)$ . Représenter la fonction de répartition des variables aléatoires suivantes :  $X$  suit la loi uniforme sur  $\{1, \dots, 5\}$  puis  $X$  suit la loi  $\mathcal{B}(5, 0.2)$ .

**TD du 7 juin** Le dernier TD auquel j'ai pu participé traité seulement des variables aléatoires.

**Exercice 1.** Une urne contient  $p$  jetons numérotés de 1 à  $p$  ( $p \geq 2$ ), on les choisit au hasard un à un et avec remise,  $X_i$  est la variable aléatoire qui donne le numéro du jeton obtenu au  $i^e$  tirage, on pose si  $n \in \mathbb{N}^* : S_n = X_1 + \dots + X_n$ .

Donner  $\mathbb{E}(S)$  et  $\text{Var}(S)$ .

**Exercice 2.** A un péage autoroutier  $n$  voitures franchissent au hasard et de manière indépendante l'une des trois barrières de péage mises à leur disposition. On note  $A, B, C$  les variables aléatoires dénombrant les voitures ayant franchi ces barrières.

1. Déterminer la loi de  $A$ .
2. Calculer les variances de  $A$ , de  $B$  et de  $A + B$ .
3. En déduire la covariance de  $A$  et  $B$ .

**Exercice 3.** On considère  $n \in \mathbb{N}^*$  et  $a_n > 0$ , on note  $X$  la variable aléatoire à valeurs dans  $\{0, \dots, n\}$  tel que si  $k \in \{0, \dots, n\}$ ,  $\mathbb{P}(X = k) = \frac{a_n}{k+1} \binom{n}{k}$ .

1. Calculer  $a_n$  pour que  $X$  soit bien une variable aléatoire.
2. Calculer  $\mathbb{E}(X)$  et  $\mathbb{E}(3X - 5)$ .

**Exercice 4.** Si  $X$  suit une loi binomiale de paramètre  $n \in \mathbb{N}^*$  et  $p \in [0; 1]$ .

1. Calculer  $\mathbb{E}(Y)$  avec  $Y = \frac{1}{X+1}$ .
2. Si  $p = 0.5$  et  $a > 0$ , calculer  $\mathbb{E}(Z)$  avec  $Z = \frac{a^X}{2n}$

**Exercice 5.** Soient  $X$  et  $Y$  deux variables aléatoires indépendantes à valeurs dans  $\{1, \dots, n\}$ .

1. Exprimer  $\mathbb{E}(X)$  en fonction de  $\mathbb{P}(X \geq k)$ .
2. On suppose que les variables  $X$  et  $Y$  suivent la loi uniforme.
  - (a) Déterminer l'espérance de  $\min(X, Y)$  puis de  $\max(X, Y)$ .
  - (b) Déterminer l'espérance de  $|X - Y|$ .

**Exercice 6.** Pour  $n \in \mathbb{N}^*$  et  $0 < p < 1$  : si  $X$  suit une loi binomiale de paramètre  $n$  et  $p$ , et  $Y = \frac{X}{n}$ .

Montrer que  $\forall \varepsilon > 0, \mathbb{P}(|Y_n - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}$

**Exercice 7.** Une secrétaire effectue, une première fois, un appel téléphonique vers  $n$  correspondants distincts. On admet que les  $n$  appels constituent  $n$  expériences indépendantes et que, pour chaque appel, la probabilité d'obtenir le correspondant demandé est  $p \in ]0; 1[$ . Soit  $X$  la variable aléatoire représentant le nombre de correspondants obtenus.

1. Donner la loi de  $X$ .
2. La secrétaire rappelle une seconde fois, dans les mêmes conditions, chacun des  $n - X$  correspondants qu'elle n'a pas pu joindre au cours de la première série d'appels. On note  $Y$  la variable aléatoire représentant le nombre de personnes jointes au cours de la seconde série d'appels.
  - (a) Soit  $i \in \llbracket 0; n \rrbracket$ . Déterminer, pour  $k \in \mathbb{N}$ ,  $\mathbb{P}_{(X=i)}(Y = k)$ .
  - (b) Prouver que  $Z = X + Y$  suit une loi binomiale dont on déterminera les paramètres.
  - (c) Déterminer l'espérance et la variance de  $Z$ .

DM

En plus des interrogations écrites, j'ai pu corriger quelques copies de *DM*, de manière à corriger les raisonnements et non juste contrôler le cours ou la réalisation d'un exercice simple. Le DM que j'ai pu corriger est fourni en annexe figure 13. J'ai donc préparé en amont le DM, et j'ai ensuite confronté mes résultats au corrigé du professeur et enfin j'ai pu corriger les copies suivant sa notation : 1 point en moins si la question est fausse, 2 points en moins si la question n'est pas traitée.

### 3.2.3. Colles

La première semaine, j'ai pu d'abord observer le fonctionnement de la colle selon l'encadrant, et ce qu'il attendait de ces élèves. L'objectif était de me préparer pour que je puisse encadrer un élève sur les trois, lors des prochaines semaines.

La première semaine où j'ai donc pu encadrer deux élèves au cours de deux colles. Le programme de colle n°28, présenté Figure 14, repose sur la fin du cours de probabilités et le début de cours sur les Matrices.

Etudiant A : Pour la première partie de la colle, il a été demandé de démontrer que  $\mathcal{T}_n^s(\mathbb{K})$  était un sous-anneau, de calculer les éléments diagonaux de  $AB$  pour  $A$  et  $B$  deux matrices triangulaires supérieurs, et finalement de montrer que  ${}^t(A \times B) = {}^tB \times {}^tA$  et  ${}^t(A^{-1}) = ({}^tA)^{-1}$ . Cette partie a été parfaitement réalisée par l'étudiant. Il a été demandé en question de cours de définir le noyau et l'image d'une matrice. J'ai ensuite donné cet exercice :

**Exercice.** Soient  $(x_n)$ ,  $(y_n)$ ,  $(z_n)$  trois suites définies par  $x_0, y_0, z_0$  et pour tout  $n \in \mathbb{N}$  :

$$\begin{cases} x_{n+1} = \frac{2}{3}x_n + \frac{1}{6}y_n + \frac{1}{6}z_n \\ y_{n+1} = \frac{1}{6}x_n + \frac{2}{3}y_n + \frac{1}{6}z_n \\ z_{n+1} = \frac{1}{6}x_n + \frac{1}{6}y_n + \frac{2}{3}z_n \end{cases}$$

Déterminer la convergence de ces 3 suites.

J'ai décidé de donner cet exercice en premier car le DM présenté Figure 13 et préparé en classe la semaine précédente pouvait donner une indication. Néanmoins, je n'attendais pas de l'étudiant qu'il sache dès le début comment résoudre cet exercice, mais le cours s'étant très bien passé, je voulais voir comme l'étudiant se débrouillerais. Il a d'abord été destabilisé par l'exercice, exotique, par rapport à



ce qui été étudié en cours, et je lui ai donc suggéré de se remémorer son DM, et de chercher une autre suite de la forme  $X_{n+1} = M \times X_n$  où  $M \in \mathcal{M}_3(\mathbb{R})$  et  $Y_n \in \mathcal{M}_{3,1}(\mathbb{R})$ . L'étudiant a alors su réagir, d'abord hésitant. Il a su trouver la matrice  $M$  demandée et calculer ses puissances sans difficultés. L'exercice a pu être résolu, peut être lentement à cause de l'hésitation. En fin d'heure, et pour utiliser les 10 dernières minutes, j'ai donné à l'étudiant :

**Exercice.** Soit  $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  défini par  $g(x, y, z) = (x - 2y, 3y + 2z)$  et les bases canoniques  $b = (e_1, e_2, e_3)$  de  $\mathbb{R}^3$  et  $B = (f_1, f_2)$  de  $\mathbb{R}^2$ .  
On considère  $e'_1 = (2, 3, 0)$ ,  $e'_2 = (1, -3, 0)$ ,  $e'_3 = (0, 1, -1)$ ,  $f'_1 = (1, 1)$ ,  $f'_2 = (1, -1)$ . On pose  $b' = (e'_1, e'_2, e'_3)$  et  $B' = (f'_1, f'_2)$ . On admet que  $b'$  et  $B'$  sont des bases respectives de  $\mathbb{R}^3$ ,  $\mathbb{R}^2$ .  
Donner les matrices de passage de  $b$  à  $b'$  et de  $B$  à  $B'$ .

Cet exercice m'a permis de contrôler que l'étudiant connaissait la définition des matrices de passage et savait les calculer. Cet exercice a été résolu mais encore avec hésitation.

Finalement, j'ai décidé de donner la note de 18 à l'étudiant, récompensant ainsi ses initiatives et ses bonnes connaissances du cours, tout en pénalisant légèrement le manque de rapidité. Néanmoins, il aurait été préférable de donner les exercices dans l'ordre inverse, de sorte à ne pas déstabiliser l'étudiant.

Etudiant B : Pour le deuxième étudiant, la même démonstration a été demandé. Il n'a pu montrer seul que le premier point, et ne savait plus ce qu'était la transposée d'une matrice, l'empêchant de poursuivre. Après avoir passé trop de temps sur le cours, j'ai demandé de définir le noyau et l'image d'une matrice, ce qui a bien été restitué, puis de passer à l'exercice :

**Exercice.** Soit  $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  défini par  $g(x, y, z) = (x - 2y, 3y + 2z)$  et les bases canoniques  $b = (e_1, e_2, e_3)$  de  $\mathbb{R}^3$  et  $B = (f_1, f_2)$  de  $\mathbb{R}^2$ .

1. Montrer que  $g$  est linéaire et donner sa matrice dans les bases  $b$ , et  $B$ .
2. On considère  $e'_1 = (2, 3, 0)$ ,  $e'_2 = (1, -3, 0)$ ,  $e'_3 = (0, 1, -1)$ ,  $f'_1 = (1, 1)$ ,  $f'_2 = (1, -1)$ . On pose  $b' = (e'_1, e'_2, e'_3)$  et  $B' = (f'_1, f'_2)$ .
  - (a) Montrer que  $b'$  et  $B'$  sont des bases respectives de  $\mathbb{R}^3$ ,  $\mathbb{R}^2$ .
  - (b) Donner les matrices de passage de  $b$  à  $b'$  et de  $B$  à  $B'$
  - (c) Donner la matrice de  $g$  dans la base  $b'$  au départ et  $B'$  à l'arrivée.

L'étudiant étant d'ores et déjà en difficulté sur le cours, j'ai donné un exercice contrôlant la connaissance de cours sur les matrices de passages, avec des questions élémentaires pour faciliter la résolution. L'étudiant n'a eu le temps de traiter seulement cet exercice

Finalement, j'ai décidé de donner la note de 13. L'étudiant faisait preuve de bonne volonté pour réussir, connaissait le cours en partie et a su réaliser l'exercice donné, ce que j'ai voulu récompenser. Néanmoins, le fait de ne pas connaître ce qu'est la transposée d'une matrice, et le fait d'avoir traité qu'un seul exercice a réduit la note. Après consultation avec l'encadrant, il était en accord avec la note décidée.

Pour la semaine suivante de colle, le programme de colle, donné Figure 15, portait intégralement sur les Matrices. J'ai eu de nouveau l'occasion d'évaluer deux étudiants, dont un d'une classe de MPSI différente de celle de l'encadrant. Le programme de colle de cette deuxième classe (MPSI1) portait sur le même chapitre, mais ne comporte pas de démonstrations choisies en amont. Les étudiants de cette classe peuvent tomber sur n'importe quelle preuve de leur chapitre.

Etudiant C : Cet étudiant (MPSI2) devait montrer que pour  $p \in \mathcal{L}(E)$  un projecteur,  $rg(p) = Tr(p)$ ,

avec  $E$  un espace vectoriel de dimension finie  $n \in \mathbb{N}^*$ . L'étudiant n'a pas su retrouver ce qu'il fallait utiliser dans les propriétés des projecteurs pour faire la démonstration, j'ai d'abord voulu orienter sa recherche en lui suggérant de réécrire  $E$  comme une somme directe judicieusement choisie. L'étudiant n'a pas su retrouver que  $\text{Ker}(p) \oplus \text{Im}(p) = E$ , ce que j'ai donc eu à rappeler. L'étudiant a alors terminé la preuve, en omettant pourquoi, pour  $B$  une base de  $\text{Im}$ ,  $\mathcal{M}_B(p) = I_r$  où  $r = \text{rg}(p)$ . J'ai alors demandé de préciser ce point, ce qui n'a pas su être donné. J'ai ensuite demandé en question de cours de donner les formules de changement de bases et de définir un système de Cramer, seul le deuxième point a pu être donné.

Le cours n'était donc que partiellement maîtrisé, il était flagrant que l'étudiant n'avait pas fourni le travail demandé ; il était néanmoins très réactif une fois les points de cours manquant donné. J'ai ensuite donné comme exercice :

**Exercice.** Soit  $A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$ , on définit :

$$u_1 = (1, 1, 1), u_2 = (1, -1, 0), u_3 = (1, 0, -1)$$

1. Montrer que  $b = (u_1, u_2, u_3)$  est une base de  $\mathbb{R}^3$ .
2. Calculer  $\text{Mat}_b(f)$  en utilisant les Formules de changement de bases.

L'étudiant a su réaliser cet exercice avec un bon rythme et sans difficulté. J'ai donc donné comme second exercice :

**Exercice.** Soit  $\text{Varphi} : \mathcal{M}_n(\mathbb{K}) \rightarrow \mathbb{K}$ , forme linéaire tel que :

$$\forall A, B \in \mathcal{M}_n(\mathbb{K}), \text{Varphi}(AB) = \text{Varphi}(BA)$$

1. Calculer  $\text{Varphi}(E_{ij})$  pour  $i, j$  différents, et montrer que  $\text{Varphi}(E_{ii})$  est constant pour tout  $i$ .
2. Montrer alors que  $\text{Varphi} = \alpha \times \text{Tr}$ .

L'étudiant a été de nouveau bloqué à cause du manque de connaissance sur le cours : en voyant que l'étudiant bloquait à la première question, j'ai suggéré de rappeler ce que valait  $E_{ip}E_{kj}$ , il n'a pas su répondre rapidement. Une fois cela rappelé, la question a été bien faite. La deuxième question a également été faite correctement.

Finalement, l'étudiant a su réaliser les exercices donné en étant très réactif au indication, et sans difficulté, dès lors qu'une connaissance du cours n'était pas nécessaire. J'ai donc donné la note de 14, pour pénaliser le manque de sérieux sur le cours.

Etudiant D : Il a été demandé à l'étudiant, venant de MPSI1, de démontrer que pour toute matrice  $A \in \mathcal{M}_{p,n}(\mathbb{K})$  de rang  $r$ , on peut extraire une matrice  $A' \in \mathcal{M}_r(\mathbb{K})$  inversible. L'étudiant a traité le cas particulier où  $A'$  est situé en haut à gauche dans la matrice  $A$ . Ce cas été bien traité, mais il avait des difficultés à voir comment passer au cas général. J'ai alors fais la remarque que, plutôt que d'écrire, pour  $C_1, \dots, C_n$  les colonnes de  $A$ ,  $\text{Vect}(C_1, \dots, C_n) = \text{Vect}(C_1, \dots, C_r)$ , il fallait obtenir  $\text{Vect}(C_1, \dots, C_n) = \text{Vect}(C_{i_1}, \dots, C_{i_r})$  avec  $(i_1, \dots, i_r) \in [[1; r]]$  distincts. L'étudiant a su reconnaître qu'il fallait utilisé le Théorème de la base extraite. Il a été demandé en question de cours les formules de changements de bases, ce qui a été restitué correctement.

L'exercice donné a été le même que l'étudiant précédent :

**Exercice.** Soit  $A = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$ , on définit :

$$u_1 = (1, 1, 1), u_2 = (1, -1, 0), u_3 = (1, 0, -1)$$

1. Montrer que  $b = (u_1, u_2, u_3)$  est une base de  $\mathbb{R}^3$ .
2. Calculer  $Mat_b(f)$  en utilisant les Formules de changement de bases.

Il a été déstabilisé par la méthode imposée, et a tenté de négocier. Je l'ai alors rassuré en disant qu'il s'agissait d'une simple application du cours et que, bien que jamais fais en exemple en cours, il pouvait s'appuyer sur les notations et les remplacer doucement. Il a alors su le faire, mais lentement.

J'ai finalement décidé de donner la note de 15 pour récompenser la bonne connaissance du cours, mais en pénalisant la lenteur sur l'exercice et le fait qu'un seul exercice ait été traité.

Pour la dernière semaine de colles, le programme de colle, donné Figure 16, reposait sur le cours sur les variables aléatoires. J'ai pu de nouveau évaluer deux étudiants de la classe.

Etudiant E : Il a été demandé de donner et montrer les espérances et variances des lois usuelles. L'étudiant a été capable de restituer la colle sans mentionner le nom de chaque loi mais seulement sa définition. J'ai donc demandé chaque nom, que l'étudiant a eu du mal à restituer. Le professeur m'a alors signalé que l'étudiant avait des difficultés et de ne faire qu'un exercice simple, j'ai donc décidé de construire, tout le long de la colle, un exercice reposant sur un couple de variables aléatoires dont les probabilités étaient données dans un tableau :

**Exercice.** Soit  $X, Y$  deux variables aléatoires telles que  $\mathbb{P}((X, Y) = (i, j))$  suit le tableau suivant pour  $i \in \{1, 2\}$  et  $j \in \{0, 1, 2\}$  :

| $i \setminus j$ | 0   | 1   | 2   |
|-----------------|-----|-----|-----|
| 1               | 0.2 | 0.3 | 0.2 |
| 2               | 0.1 | 0   | 0.2 |

1. Vérifier que  $(X, Y)$  est bien une variable aléatoire.
2. Déterminer les lois marginales de  $X$  et  $Y$ .
3. Déterminer la loi de  $Z = XY$ .
4. Calculer la loi de  $Y$  sachant  $X = 1$  puis  $X = 2$ .

L'étudiant a réussi chacune de ces questions avec de l'aide pour la dernière. L'exercice a été fait lentement, mais après indication de l'encadrant, il a été décidé de mettre la note de 14.

Etudiant F : Il a été demandé de montrer que  $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$  pour deux variables aléatoires indépendantes, puis de calculer  $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$ . La démonstration a été réalisée parfaitement et rapidement. Il a été donné en exercice les deux exercices suivants, dans l'ordre :

**Exercice.** Soit  $X_1, \dots, X_n$  sont des variables aléatoires de même espérance, de même covariance deux à deux, i.e. pour tout  $i, j \in \{1, \dots, b\}$ ,  $Cov(X_i, X_j) = \alpha$ . Calculer  $\mathbb{E}(S)$  et  $Var(S)$  où  $S = X_1 + \dots + X_n$ .

**Exercice.** On considère  $(X_n)_{n \geq 1}$  une suite de variables aléatoires mutuellement indépendantes suivant toutes la loi  $\mathcal{B}(p)$  où  $0 < p < 1$  et  $S_n = X_1 + \dots + X_n$  et  $u_n$  la probabilité pour que  $S_n$  pour  $n \in \mathbb{N}^*$ .

1. Déterminer la loi de  $S_n$ , pour  $n \in \mathbb{N}^*$ .
2. Calculer  $u_1$ ,  $u_2$  et  $u_3$ .
3. Montrer qu'il existe  $a, b \in \mathbb{R}$  tels que  $\forall n \in \mathbb{N}^* : au_{n+1} = au_n + b$ .
4. En déduire l'expression et la limite de  $u_n$ .

Les exercices ont été parfaitement réalisés, la note de 19 a donc été donné.

### 3.2.4. Correction d'exercices et démonstrations

Enfin, une autre tâche confiée était la réalisation au tableau de démonstration et de correction d'exercices que j'avais préalablement préparée. La correction d'exercice intervenait principalement en début d'heure et concernait des exercices d'applications élémentaires du cours. Le premier chapitre abordé lors de mon stage était celui sur les matrices. Malheureusement, les examens ont été déplacés pendant mon stage, m'empêchant de faire beaucoup sur ce chapitre. J'ai tout de même eu l'occasion de corriger deux exercices :

**Exercice.** On considère  $\varphi : \begin{cases} \mathbb{R}_2[X] \rightarrow \mathbb{R}_2[X] \\ P \mapsto P + P' \end{cases}$ , on pose  $B = (1, 1 + X, 1 + X + X^2)$  une base de  $\mathbb{R}_2[X]$ .

1. Ecrire la matrice de  $\varphi$  dans la base canonique  $B_0$  de  $\mathbb{R}_2[X]$ .
2. Ecrire la matrice de  $\varphi$  dans la base  $B$  de  $\mathbb{R}_2[X]$ .

**Exercice.** 1. Déterminer les matrices équivalentes à  $0_{p,n}$ , puis à  $I_n$ .

2. Les deux matrices suivantes sont-elles équivalentes ?  $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$  et  $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ .

La semaine suivante reposait sur les variables aléatoires. J'ai assisté à l'intégralité du chapitre et j'ai donc pu préparer deux parties de cours, la première sur les espérances et variances des lois usuelles, la seconde sur les inégalités de Markov et Bienaymé-Tchebychev, comme présenté page 49. J'ai également pu faire, au tableau, la correction des exercices suivants :

**Exercice.** Soit  $X$  une variable aléatoire réelle sur  $\Omega$  fini à valeurs dans  $\{0, \dots, N\}$ , montrer que  $\mathbb{E}(X) = \sum_{k=0}^{N-1} \mathbb{P}(X > k)$

**Exercice.** Un joueur tire  $n$  boules successivement avec remise dans une urne contenant  $N$  boules numérotées de 1 à  $N$ ,  $X$  (resp.  $Y$ ) est la variable aléatoire qui donne le numéro de la plus grand (resp. petite) boule tirée.

Donner la loi de  $X$  et la loi de  $Y$ .

## B. Espérances et Variances des lois usuelles

## B.1. Définition et espérances des lois usuelles

Définition 3.2.1 -

Soit  $X$  une variable aléatoire réelle sur  $(\Omega, \mathbb{P})$ . Si  $X(\Omega) = \{x_1, \dots, x_n\}$  alors on appelle **espérance de  $X$**  le réel noté  $\mathbb{E}(X)$  défini par :

$$\mathbb{E}(X) = \sum_{k=1}^n x_k \mathbb{P}(X = x_k)$$

Proposition 3.2.0 -

Soit  $X$  une variable aléatoire sur  $(\Omega, \mathbb{P})$ .

- i. Si  $X$  est constante égale à  $a$ , alors  $\mathbb{E}(X) = a$
- ii. Si  $X$  suit la loi uniforme avec  $X(\Omega) = \{a_1, \dots, a_n\} \subset \mathbb{R}$ , alors  $X$  est d'espérance finie avec  $\mathbb{E}(X) = \frac{1}{n} \sum_{k=1}^n a_k$ .
- iii. Si  $X$  suit la loi de Bernoulli de paramètre  $p \in [0; 1]$ , alors  $X$  est d'espérance finie et  $\mathbb{E}(X) = p$ .
- iv. Si  $X$  suit la loi binomiale de paramètre  $n \in \mathbb{N}^*$  et  $p \in [0; 1]$ , alors  $X$  est d'espérance finie et  $\mathbb{E}(X) = np$ .

*Démonstration.* i.  $X(\Omega) = \{a\}$  et  $\mathbb{P}(X = a) = 1$ , donc  $\mathbb{E}(X) = a \times \mathbb{P}(X = a) = a$ .

ii. Pour tout  $k \in \{1, \dots, n\}$ ,  $\mathbb{P}(X = a_k) = \frac{1}{n}$ , donc  $\mathbb{E}(X) = \sum_{k=1}^n a_k \mathbb{P}(X = a_k) = \sum_{k=1}^n \frac{a_k}{n} = \frac{1}{n} \sum_{k=1}^n a_k$ .

iii. Si  $X$  suit la loi de Bernoulli de paramètre  $p \in [0, 1]$ , on a  $X(\Omega) = \{0; 1\}$  :

$$\mathbb{E}(X) = 0 \times \mathbb{P}(X = 0) + 1 \times \mathbb{P}(X = 1) = 0 \times (1 - p) + 1 \times p = p$$

iv. Si  $X$  suit la loi binomiale de paramètre  $n \in \mathbb{N}^*$  et  $p \in [0; 1]$ , on a  $X(\Omega) = \{0, \dots, n\}$  :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^n k \mathbb{P}(X = k) = \sum_{k=0}^n \binom{n}{k} k p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n \binom{n}{k} k p^k (1-p)^{n-k} \end{aligned}$$

Si  $n \geq 1 : k \in \{1, \dots, n\} : k \binom{n}{k} = n \binom{n-1}{k-1}$ . Donc :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=1}^n n \binom{n-1}{k-1} p^k (1-p)^{n-k} = n \sum_{j=0}^{n-1} \binom{n-1}{j} p^{j+1} (1-p)^{n-1-j} \\ &= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} = np \end{aligned}$$

□

*Remarque.* Si  $X$  suit la loi uniforme sur  $\{1, \dots, n\}$ , alors  $\mathbb{E}(X) = \frac{n+1}{2}$ .

## B.2. Définition et variances des lois usuelles

## Définition 3.2.2 -

Soit  $X$  une variable aléatoire réelle sur  $(\Omega, \mathbb{P})$ , on appelle variance de  $X$  le réel positif  $\text{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$ .

## Proposition 3.2.0 -

Soit  $X$  une variable aléatoire sur  $(\Omega, \mathbb{P})$ .

- i. Si  $X$  est constante alors sa variance est nulle.
- ii. Si  $X$  suit la loi uniforme sur  $\{1, \dots, n\}$ , alors  $\text{Var}(X) = \frac{n^2 - 1}{12}$ .
- iii. Si  $X$  suit la loi de Bernoulli de paramètre  $p \in [0; 1]$ , alors  $\text{Var}(X) = p(1-p)$ .
- iv. Si  $X$  suit la loi binomiale de paramètre  $n \in \mathbb{N}^*$ ,  $p \in [0; 1]$ , alors  $\text{Var}(X) = np(1-p)$ .

*Remarque.* La formule de transfert ainsi que celle de Koenig-Huygens ont été énoncés et montrés par l'encadrant.

*Démonstration.* i.  $X(\Omega) = \{a\}$ ,  $\mathbb{P}(X = a) = 1$  donc  $\mathbb{E}(X^2) = a^2$  et  $\mathbb{E}(X) = a$ .

Finalement,  $\text{Var}(X) = a^2 - a^2 = 0$  d'après la formule de Koenig-Huygens.

ii. Pour tout  $k \in \{1, \dots, n\}$ ,  $\mathbb{P}(X = k) = \frac{1}{n}$ , donc :

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=1}^n k \frac{1}{n} = \frac{n+1}{2} \\ \mathbb{E}(X^2) &= \sum_{k=1}^n k^2 \frac{1}{n} = \frac{(n+1)(2n+1)}{6}\end{aligned}$$

et donc on a :

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2 - 1}{12}$$

iii. Si  $X$  suit la loi de Bernoulli de paramètre  $p \in [0; 1]$ , on a :

$$\begin{aligned}\mathbb{E}(X) &= p \\ \mathbb{E}(X^2) &= 0^2 \times (1-p) + 1^2 \times p = p\end{aligned}$$

Donc  $\text{Var}(X) = p - p^2 = p(1-p)$

iv. Si  $X$  suit la loi binomiale de paramètre  $n \geq 2$  et  $p \in [0; 1]$ , on a :

$$\begin{aligned}\mathbb{E}(X) &= np \\ \mathbb{E}(X^2) &= \mathbb{E}(X^2 - X + X) = \mathbb{E}(X(X-1)) + np\end{aligned}$$

et on remarque que :

$$\begin{aligned}\mathbb{E}(X(X-1)) &= \sum_{k=0}^n k(k-1) \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=2}^n k(k-1) \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=2}^n n(n-1) \binom{n-2}{k-2} p^k (1-p)^{n-k} \\ &= n(n-1)p^2 \sum_{j=0}^{n-2} \binom{n-2}{j} p^j (1-p)^{n-2-j} = n(n-1)p^2\end{aligned}$$

Finalement,  $\text{Var}(X) = n(n-1)p^2 + np - (np)^2 = np(np - p + 1 - np) = np(1-p)$ .

*Remarque.* Si  $n = 1$ , alors  $\text{Var}(X) = 1 \times p(1 - p) = np(1 - p)$ .

□

### 3. Estimation de la dispersion

#### Théorème 3.2.1 - Inégalité de Markov

Soit  $X$  une variable aléatoire réelle sur  $(\Omega, \mathbb{P})$ , positive, alors, pour tout  $\varepsilon > 0$  :

$$\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}(X)}{\varepsilon}$$

*Démonstration.* Soit  $\varepsilon > 0$  fixé. On pose  $X(\Omega) = \{x_1, \dots, x_k\}$ .

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=1}^n x_k \mathbb{P}(X = x_k) = \sum_{x_k \geq \varepsilon} x_k \mathbb{P}(X = x_k) + \sum_{x_k < \varepsilon} \underbrace{x_k}_{\geq 0} \mathbb{P}(X = x_k) \\ &\geq \sum_{x_k \geq \varepsilon} \underbrace{x_k}_{\geq \varepsilon} \mathbb{P}(X = x_k) \\ &\geq \varepsilon \sum_{x_k \geq \varepsilon} \mathbb{P}(X = x_k) = \varepsilon \mathbb{P}(X \geq \varepsilon) \end{aligned}$$

Donc  $\mathbb{E}(X) \geq \varepsilon \mathbb{P}(X \geq \varepsilon)$ . On retrouve bien l'inégalité de Markov.

□

#### Théorème 3.2.2 - Inégalité de Bienaymé-Tchebychev

Soit  $X$  une variable aléatoire réelle sur  $(\Omega, \mathbb{P})$ . Alors, pour tout  $a > 0$ ,  $\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\text{Var}(X)}{a^2}$

*Démonstration.* Soit  $a > 0$  fixé. En posant  $Y = (X - \mathbb{E}(X))^2$ , on a  $Y$  une variable aléatoire positive. De plus, on remarque que  $(|X - \mathbb{E}(X)| \geq a) = (Y \geq a^2)$ .

On a donc d'après l'inégalité de Markov appliqué à  $Y$  :

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) = \mathbb{P}(Y \geq a^2) \leq \frac{\mathbb{E}(Y)}{a^2}$$

Et  $\mathbb{E}(Y) = \text{Var}(X)$ . On retrouve bien l'inégalité de Bienaymé-Tchebychev.

□

Les dernières semaines furent consacrées principalement au chapitre sur le déterminant. J'ai pu aller de nombreuses fois au tableau, et présenter des propriétés sur le déterminant ainsi qu'une démonstration du calcul du déterminant de Vandermonde, présenté à la page 52. J'ai pu faire la correction des exercices suivants :

**Exercice.** On pose pour  $n \in \mathbb{N}^*$ , le déterminant  $n \times n$  :

$$D_n = \begin{vmatrix} 10 & 3 & & (0) \\ 3 & \ddots & \ddots & \\ & \ddots & \ddots & 3 \\ (0) & & 3 & 10 \end{vmatrix}$$

**Exercice.** Calculer les déterminants  $n \times n$  suivants :

$$\text{i. } D_1 = \begin{vmatrix} 0 & \dots & 0 & \lambda_n \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \lambda_1 & 0 & \dots & 0 \end{vmatrix}$$

$$\text{ii. } D_2 = \begin{vmatrix} a & a_{12} & \dots & a_{1n} \\ a & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a & \dots & a & a \end{vmatrix}$$

**Exercice.** Soit  $f : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \rightarrow \mathcal{M}_n(\mathbb{R}) \\ A & \mapsto {}^t A \end{cases}$

1. Montrer que  $f$  est une application linéaire.
2. Si  $n = 2$ , calculer  $M$  la matrice de  $f$  dans la base canonique de  $\mathcal{M}_n(\mathbb{R})$  et  $\det(f)$ .
3. Calculer  $f \circ f$ . Que dire de  $\det(f)$  ?
4. Calculer  $\det(f)$  en utilisant  $\mathcal{M}_n(\mathbb{R}) = \mathcal{S}_n(\mathbb{R}) \oplus \mathcal{A}_n(\mathbb{R})$ .

B. Propriétés du déterminant d'endomorphisme :

Proposition 3.2.2 -

Si  $E$  est un  $\mathbb{K}$ -espace vectoriel de dimension finie  $n \in \mathbb{N}^*$ .

- i.  $\det(\text{Id}_E) = 1$
- ii. Si  $u, v \in \mathcal{L}(E)$ , alors  $\det(uov) = \det(vou) = \det(u) \det(v)$
- iii. Si  $\alpha \in \mathbb{K}^*$ , et  $u \in \mathcal{L}(E)$  :  $\det(\alpha u) = \alpha^n \det(u)$

*Remarque.* Il a été montré précédemment par leur professeur que pour  $u$  un endomorphisme de  $E$  de dimension finie, de base  $b$ . Pour toutes familles de vecteurs  $(x_1, \dots, x_n) \in E^n$  :  $\det_b(u(x_1), \dots, u(x_n)) = \det(u) \det_b(x_1, \dots, x_n)$

*Démonstration.* Soit  $b = (e_1, \dots, e_n)$  une base de  $E$ .

- i.  $\det(\text{Id}_E) = \det_b(\text{Id}_E(b)) = \det_b(e_1, \dots, e_n) = 1$  par définition du déterminant.
- ii. On a  $\det(uov) = \det_b(u(v(e_1)), \dots, u(v(e_n))) = \det(u) \det_b(v(e_1), \dots, v(e_n))$ , soit  $\det(uov) = \det(u) \det(v) \det_b(e_1, \dots, e_n) = \det(u) \det(v)$ .  
On a donc finalement  $\det(uov) = \det(u) \det(v) = \det(v) \det(u) = \det(vou)$ .



iii. Si  $\alpha \in \mathbb{K}^*$ ,  $\det(\alpha u) = \det(\alpha u(e_1), \dots, \alpha u(e_n)) = \alpha^n \det(u(e_1), \dots, u(e_n))$  par  $n$ -linéarité, ce qui donne :

$$\det(\alpha u) = \alpha^n \det(u)$$

□

### B. Matrices de Vandermonde :

#### Définition 3.2.3 -

On appelle matrice de Vandermonde une matrice du type

$$A = \begin{pmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{pmatrix}$$

avec  $a_1, \dots, a_n \in \mathbb{K}$ .

**Exemple 3.2.1.** Pour  $n = 2$  :  $\det(A) = \begin{vmatrix} 1 & a \\ 1 & b \end{vmatrix} = (b - a)$

#### Proposition 3.2.2 -

Si  $a_1, \dots, a_n \in \mathbb{K}$ , alors

$$\det \begin{pmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{pmatrix} = \prod_{i < j} (a_j - a_i)$$

*Démonstration.* On montre la récurrence sur  $n \in \mathbb{N}$  avec  $n \geq 2$ .

- Initialisation Le cas  $n = 2$  a été traité en exemple.
- Hérédité On suppose avoir établi le résultat pour un certain  $n - 1 \in \mathbb{N}$  avec  $n - 1 \geq 2$ , on considère alors :

$$A_n = \begin{pmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

avec  $a_1, \dots, a_n \in \mathbb{K}$

On considère  $P(X) = (X - a_1) \dots (X - a_{n-1})$ ,  $P$  est un polynôme de degré  $n - 1$  qui a  $a_1, \dots, a_{n-1}$  comme racines. On donne sa forme développée :

$$P(X) = X^{n-1} + \alpha_{n-2}X^{n-2} + \dots + \alpha_1X + \alpha_0$$

On calcule alors :

$$\det(A_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix} \stackrel{C_n \leftarrow C_n + \alpha_{n-2}C_{n-1} + \dots + \alpha_0 C_1}{=} \begin{vmatrix} 1 & a_1 & \dots & P(a_1) \\ 1 & a_2 & \dots & P(a_2) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \dots & P(a_n) \end{vmatrix}$$

$$\det(A_n) = \begin{vmatrix} 1 & a_1 & \dots & a_1^{n-2} & 0 \\ 1 & a_2 & \dots & a_2^{n-2} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & a_{n-1} & \dots & a_{n-1}^{n-2} & 0 \\ 1 & a_n & \dots & a_n^{n-2} & P(a_n) \end{vmatrix}$$

En développant ce déterminant par rapport à la dernière colonne on obtient :

$$\det(A_n) = P(a_n) \det(A_{n-1})$$

Donc en appliquant l'hypothèse de récurrence :

$$\det(A_n) = P(a_n) \prod_{1 \leq i < j \leq n-1} (a_j - a_i) = \prod_{1 \leq i < j \leq n} (a_j - a_i)$$

Par récurrence, le résultat est vrai pour tout  $n \in \mathbb{N}$  avec  $n \geq 2$ . □

Pour terminer, j'ai eu le temps, à la fin du stage, d'assister au début du cours sur les espaces préhilbertiens et euclidiens, et j'ai donc pu présenter la démonstration de l'inégalité de Cauchy Schwarz :

Proposition 3.2.2 - Inégalité de Cauchy-Schwarz

Soit  $(E, <, >)$  un espace préhilbertien réel. Soit  $x, y \in E$  :

$$|<x, y>| \leq \sqrt{<x, x>} \sqrt{<y, y>}$$

avec égalité si et seulement si  $(x, y)$  est liée dans  $E$ .

*Démonstration.* On considère  $x, y \in E$  :

Si  $y = 0_E$  alors  $<x, y> = 0$  donc  $0 = |<x, y>| \leq \sqrt{<x, x>} \sqrt{<y, y>} = 0$ .

Si  $y \neq 0_E$ , on pose pour  $\lambda \in \mathbb{R} : T(\lambda) = <x + \lambda y, x + \lambda y>$ .

Par positivité du produit scalaire,  $T(\lambda) \geq 0$  pour tout  $\lambda \in \mathbb{R}$ , or on remarque que :

$$T(\lambda) = <x, x> + 2\lambda <x, y> + \lambda^2 <y, y>$$

Finalement,  $T$  est un trinôme du second degré en  $\lambda$  (car  $<y, y> \neq 0$ , sinon  $y = 0_E$ ) de signe constant. Donc son discriminant  $\Delta$  est négatif ou nul :

$$\Delta = 4(<x, y>^2 - <x, x> <y, y>) \leq 0$$

On retrouve finalement que  $|<x, y>| \leq \sqrt{<x, x>} \sqrt{<y, y>}$ .

Considérons maintenant le cas d'égalité :

$\Rightarrow$  Si  $|<x, y>| = \sqrt{<x, x>} \sqrt{<y, y>}$  ; pour  $y \neq 0_E$  on a alors que  $\Delta$  défini précédemment est nul. Donc le trinôme  $T$  admet une racine  $\lambda_0 \in \mathbb{R}$ . De fait,  $T(\lambda_0) = <x + \lambda_0 y, x + \lambda_0 y> = 0$ , donc par le caractère défini du produit scalaire,  $x + \lambda_0 y = 0$  et donc  $(x, y)$  est liée dans  $E$ .

Le cas  $y = 0_E$  est évident

$\Leftrightarrow$  Si  $(x, y)$  est liée dans  $E$ , alors  $x, y$  sont colinéaires, donc  $\exists \alpha \in \mathbb{R}$  tel que  $x = \alpha y$  ou  $y = \alpha x$  (sans perte de généralité, on considère  $x = \alpha y$ ).

Dans ce cas :  $|\langle x, y \rangle| = |\langle \alpha y, y \rangle| = |\alpha| \langle y, y \rangle$

et  $\sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle} = \sqrt{\alpha^2 \langle y, y \rangle^2} = |\alpha| \langle y, y \rangle = |\langle x, y \rangle|$  □

---

## Conclusion

---

En conclusion, ce stage m'a permis de m'approcher au mieux du métier de professeur. J'ai pu découvrir autant les compétences écrites qu'orales nécessaires pour exercer ce métier. La multitude de tâches confiées et l'autonomie offerte par mon encadrant ont été une chance.

J'ai appris, grâce à ce stage, à préparer des TD, des colles et des exercices et de me préparer à adapter ces derniers aux étudiants devant moi. Le fonctionnement des TD m'a permis d'expliquer à de multiples reprises et de plusieurs manières différentes les exercices, de manière à s'accorder au mieux au profil de l'étudiant en face. J'ai pu préparer des parties de cours, brève pour ne pas prendre du retard sur le programme, mais suffisante pour être confronté à une situation conforme à la réalité d'un professeur. Avoir pu être autant présente lors des TD, colles et cours, m'a permis également d'établir un lien avec les étudiants, qui, en conséquence, venait à ma rencontre pour poser des questions sur d'autres exercices fournis, sur le DM à préparer ou même pour des questions d'orientation et de choix de spé.

## Deuxième Partie : Formation

,map size=0.9,widthh=

# Lars Hörmander et The analysis of Linear Partial Differential Operators



Auriane GABAUT  
Juin 2022

## Sommaire

|                                                                 |    |
|-----------------------------------------------------------------|----|
| Introduction et Motivations . . . . .                           | 57 |
| 4.1 Hörmander et ses travaux avant la médaille Fields . . . . . | 58 |
| 4.2 Un apport conséquent dans le domaine des EDP . . . . .      | 60 |
| 4.3 Reconnaissance et importance de son travail . . . . .       | 62 |
| Conclusion . . . . .                                            | 63 |

## Introduction et Motivations

Lars Hörmander est un mathématicien suédois né le 24 janvier 1931 et décédé le 25 novembre 2012. Il a principalement travaillé sur les Équations aux Dérivées Partielles (EDP). Il a été l'un des mathématiciens les plus influents du XX<sup>e</sup> et a joué un rôle fondamental dans le développement de ce domaine des mathématiques. Il est en particulier reconnu pour sa capacité à synthétiser, à produire des textes concis, précis, ne regroupant que l'essentiel. Pour son travail, il a été récompensé en 1962 par la médaille Fields, le prix Wolf en 1988 et enfin le prix Steele en 2006.

Son traité en 4 volumes, et ce sur quoi ce mémoire va porter, nommé *The Analysis of Linear Partial Differential Operators* [Hö83], [Hö85], est considéré comme une référence sur le sujet. Tout au long de ce dossier, nous allons tenter de comprendre dans quelles mesures son travail sur le domaine des E.D.P. a contribué à faire avancer les mathématiques et l'impact qu'il a eu sur le domaine. Le premier volume de ce livre a été partiellement traité au cours d'un TER encadré par Patrick Gérard, ce qui est la principale motivation de ce sujet.

## 4.1. Hörmander et ses travaux avant la médaille Fields



FIGURE 4.1 – Lars Hörmander en 1969

Nous allons commencer par présenter le travail de Hörmander avant la médaille Fields, et ce qui l'a conduit à rédiger *The Analysis of Linear Partial Differential Operators* [Hö83], [Hö85]. D'abord, on peut remarquer que sa carrière de mathématicien semble commencer dès sa thèse *On the theory of general partial differential operators* [Hör55], encadré par L. Gårding. En effet, elle a été publiée en 1955 dans le magazine *Acta Mathematica* et est d'ores et déjà un document riche sur le sujet. Dans cette thèse, de 300 pages, Hörmander a travaillé presque indépendamment, il a très peu sollicité son directeur de thèse et a lui même trouvé les problèmes qu'il souhaitait traiter et résoudre.

It is no exaggeration to say that the thesis opened a new era of the subject of partial differential equation. Some of the most important results were quite original and had not even been envisioned before.

Jan Boman and Ragnar Sigurdsson, [BS15]

Une chose importante à constater est la distance prise avec Hörmander, tout au long de sa thèse, avec la théorie de Schwarz. En effet, bien que la théorie des distribution de L. Schwarz ait été développée dans les années 40', beaucoup de mathématiciens sont sceptiques sur le sujet, pourtant dorénavant presque indispensable à l'étude des EDP.

Une deuxième contribution est un article dans *Mathematica Scandinavica* [Hö58] *On the uniqueness of the Cauchy Problem* datant de 1958, où il propose une nouvelle preuve à la caractérisation de l'hypoellipticité des EDP à coefficients constants, qu'il améliore même en apportant des précisions sur la régularité des solutions. Si on définit :

$$\text{car}P = \{\xi \in \mathbb{C}^n, P(\xi) = 0\}$$

on a que  $P$  est hypo-elliptique si et seulement si :

$$|\xi| \rightarrow \infty, \forall \xi \in \text{car} P \Rightarrow |\text{Im} \xi| \rightarrow \infty$$

La preuve fournie par Hörmander est innovante, personne n'avait alors proposé un tel résultat.

Enfin, la publication de *Linear Partial Differential Operators* en 1963 [Hö63] met en lumière toute la théorie des distributions de L. Schwartz, jusqu'alors dévalorisée. A ce propos, il est dit par exemple dans une revue pour le journal *Science* [Ehr64] :

Linear Partial Differential Operators, by Hörmander, represents the progress that has been made in partial differential equations as a result of this new view- point. The author himself is a major contributor to that progress. [...] Linear Partial Differential Operators contains very deep results. It is lucid and extremely readable, and it should become one of the classics in the field.

---

Leon Ehrenpreis, [Ehr64]



FIGURE 4.2 – Lars Hörmander recevant la médaille Fields en 1962

C'est à cette période qu'il entame la rédaction du livre en 4 volumes *The analysis of Linear Partial Differential Operators* à la suite d'une demande de la maison d'édition *Springer Grundlehren*. Plus tard, Hörmander commence à travailler sur les opérateurs de Lewy et se fixe comme objectif de comprendre géométriquement une classe d'opérateur. Et c'est dans les articles *Differential operators of principal type* [Hö60] et *Estimates for translation invariant operators on  $L^p$  spaces* [Hör60] publiés tous deux en 1960 qu'il y parvient. Hörmander donne alors une généralisation des opérateurs de Lewy et de nouveaux résultats dessus. Il reçoit finalement pour son travail dans le domaine des EDP, effectué depuis la fin de sa thèse, et en particulier pour le livre [Hö63] *Linear Partial Differential Operator* la médaille Fields en 1962.

His impressive work on Partial Differential Equations, in particular his characterization of hypoellipticity for constant coefficients and his geometrical explanation of the Lewy non-solvability phenomenon were certainly very strong arguments for awarding him the Medal. Also his new point of view on Partial Differential Equations, which combined functional analysis with 'a priori' inequalities, had led to very general results on large classes of equations, which had been out of reach in the early fifties.

---

Nicolas Lerner, [Ler13]

Ainsi, dans ce domaine des mathématiques, de nombreux mathématiciens ne voyaient pas d'un bon oeil le cadre créé par L. Schwartz pour l'étude et la résolution des EDP, mais Hörmander décide d'utiliser explicitement et de présenter toute la théorie de Schwartz. Hörmander lui décide de traiter le sujet de manières approfondie et concise. Les preuves sont abrégées à l'essentiel. Le but était de donner une étude systématique des questions d'existence, d'unicité et de régularité des solutions des équations aux dérivées partielles et des problèmes aux limites. Ses différents textes et articles sont une première synthèse des résultats connus sur les la théorie des EDP et deviennent rapidement des références pour les chercheurs ; il ne s'adresse néanmoins qu'à un public restreint. En résumé, Hörmander présente les progrès réalisés dans le passé et même dans un passé proche dans ce domaine, c'est une synthèse lisible, concise et complète sur le sujet pour lequel il apporte de nouvelles propriétés jusqu'alors pas envisagées.

Le travail de Hörmander est d'ores-et-déjà une référence et après la réception de la médaille Fields, il entame son travail de rédaction pour son livre *The Analysis of Linear Partial Differential Operators* [Hö63]. Nous avons en particulier travaillé sur le premier volume, accessible à notre niveau.

---

## 4.2. Un apport conséquent dans le domaine des EDP

---

Son oeuvre en 4 volumes a été entamée les étés 1960-1961, à la suite d'une offre de la maison d'édition *Springer Grundlehren series* pour publier un livre sur les EDP. Hörmander publie alors en 1963 *Linear partial Differential operators*. Nicolas Lerner [Ler13] dit à propos de ce livre :



That book was a milestone in the study of PDE, and a large mathematical public discovered L. Hörmander's exposition of recent progress in the area.

---

Nicolas Lerner, [Ler13]

Une nouvelle fois, Hörmander met au centre la théorie des distributions développée par Schwarz, comme un outil indispensable à l'étude des EDP. Dans ce premier livre, il étudie d'abord les EDP liées à des opérateurs à coefficients constants, et à des coefficients variables.

A propos de la rédaction de ce livre, Hörmander dit :

Some time after two summers (1960, 1961) at Stanford, I received an offer of a part time appointment as professor at Stanford University[...] I had barely arrived at Stanford when I received an offer to come to the Institute for Advanced Study as permanent member and professor. Although I had previously been determined not to leave Sweden, the opportunity to do research full time in a mathematically very active environment was hard to resist[...] I decided in the fall of 1963 to accept the offer from the IAS and resign from the universities of Stockholm and Stanford to take up a new position in Princeton in the fall of 1964.

---

Lars Hörmander, [Hö83]

Sa carrière a pris un tournant important à partir de la médaille fields, et ce premier livre a incité beaucoup d'autres mathématiciens à s'intéresser au sujet et notamment sur l'opérateur :

$$\mathcal{H} = X_0 - \sum_{1 \leq j \leq r} X_j^2$$

appelé *Hörmander's sum of squares*.

Ensuite, en 1972 est publié avec Duistermaat et Hörmander *Fourier integral operators* [JD72], contenant un grand nombre de connaissances sur les Intégrales de Fourier. Ce nouveau texte offre une description complète des singularités dans les transformées de Fourier.

C'est ainsi qu'en mars 1982, Hörmander reçoit un "doctorat honoris causa" (un titre honorifique) décerné par l'université Paris-Sud à Orsay. Il avait alors entrepris depuis 1979 de travailler sur un livre sur l'ensemble des connaissances relatives aux EDP ; cet ouvrage était largement attendu par les scientifiques du domaine. Les deux premiers volumes sont publiés en 1983.

Le premier est *Distribution Theory and Fourier Analysis*. Il rappelle toute la théorie des intégrales de Fourier et des distributions et devient rapidement une référence dans le domaine de l'analyse. Il y propose également une nouvelle preuve du théorème du noyau de Schwarz, du théorème de Malgrange et une étude plus fine des phases stationnaires, avec une estimation plus précise.

Le deuxième volume *Differential Operators with Constant Coefficients* présente en préface :

This volume is an expanded version of Chapters III, IV, V and VII of my 1963 book[...] The main technical tool in this volume is the FourierLaplace transformation. More powerful methods for the study of operators with variable coefficients will be developed in Volume III. However, the constant coefficient theory has given the guidelines for all that work. Although the field is no longer very active - perhaps because of its advanced state of development -[...] the material presented here should not be neglected by the serious student who wants to get a balanced perspective of the theory[...].

---

Ces deux premiers livres sont alors d'abord vus comme une "expansion" de son premier livre publié en 1963. Il apporte de nouvelles preuves originales, et plus concises. Hörmander veut être efficace dans ses explications, et aller droit à l'essentiel, tout en restant néanmoins accessible aux étudiants aguérés.

Enfin, les deux derniers volumes de ce livre, *Pseudodifferential operators* et *Fourier integral operators*, sont publiés en 1985. Contrairement au premier volume, ces deux derniers volumes ne s'adressent qu'à des chercheurs, déjà bien armés sur le sujet. Il apporte autant d'informations sur les outils déjà mis en place que sur de nouvelles propriétés :

Since I needed this result in my research, I consulted Hörmander. He was optimistic, but the full proof did not emerge until the second volume of his magnificent four-volume opus *The Analysis of Linear Partial Differential Operators*.

---

Sigurdur Helgason,

---

### 4.3. Reconnaissance et importance de son travail

---

On a donc vu que la rédaction du livre en 4 volumes *The analysis of Linear Partial Differential Operators* entre 1983 et 1985 couvre largement le domaine des EDP. De fait et pour sa contribution importante, La Société Américaine de mathématiques lui décerne le prix Steele en 2006 pour ce livre. Il est alors dit à cette occasion et à son sujet :

In the history of mathematics, one is hard-pressed to find any comparable 'expository' work that covers so much material, and with such depth and understanding of such a broad area of mathematics.

---

De fait et d'après ses compères, Hörmander a joué un rôle déterminant dans le domaine des EDP et des intégrales de Fourier. Il a élaboré une théorie concise et précise sur les opérateurs pseudo-différentiels, et a provoqué une réelle avancée dans ce domaine. En effet, bien que déjà employée en physique, la théorie sur les intégrales de Fourier d'un point de vue mathématique est essentiellement dû au travail de Hörmander et est tardif par rapport à son emploi. Comme l'atteste Jean-Michel Bony, un mathématicien et professeur à l'école Polytechnique, et Christer Oscar Kiselman, mathématicien et professeur en université, le travail de Hörmander est une source précieuse tant pour la recherche que pour l'enseignement.

Without Lars Hörmander, this development would have been certainly much slower, and it could have produced a messy and tangled heap of results and specific methods. We are all indebted to him and will remain so for a long time. [...] Above all, it is in his book *Linear Partial Differential Operators* that I learned the theory of PDE when I was a student in École Normale Supérieure. I could appreciate then the precision and the concision of his style : a complete exposition in eighteen pages of the theory of distributions, which was at that time the subject of a one-semester course in Paris.

---

Jean-Michel Bony, [BS15]

Nobody has been so important for my intellectual and scientific development as Lars.

---

Christer O. Kiselman [BS15]

---

## Conclusion

---

En résumé, le rôle de Hörmander dans la théorie des EDP est déterminant. Il a autant permis de rendre accessible et populaire la théorie des distributions et son application aux domaines des EDP, que permis des découvertes et une avancée majeure dans le domaine. Il a par exemple été le premier à étudier l'opérateur de Hörmander et ses propriétés hypoelliptiques qui est pourtant un opérateur important maintenant en théorie des probabilités et en géométrie.

Mais on pourrait ajouter également que ses contributions dans d'autres domaines ne sont pas non plus sans conséquences comme en analyse complexe où *An introduction to Complex Analysis in Several variables*, est par exemple une référence incontournable dans le domaine. Il a été construit à partir de notes de cours qu'il donnait à Stanford en 1964.

Enfin, chacun des livres publiés par Hörmander est reconnu et utilisé tant par les étudiants que les chercheurs. Il s'agit d'ouvrages synthétiques, et pourtant complets.

---

## Liste des références

---

J. Boman and R. Sigurdsson. To the memory of lars hörmänder (1931-2012). *Notices of the American Mathematical Society*, 2015.

L. Ehrenpreis. Mathematical analysis : *linear partial differential operators*. lars hörmänder. springer, berlin ; academic press, new york, 1963. viii + 287 pp. illus. \$10.50. *Science*, 143(3603) :234–234, 1964.

Lars Hörmander. On the theory of general partial differential operators. *Acta Mathematica*, 94(none) :161 – 248, 1955.

Lars Hörmander. Estimates for translation invariant operators in  $L^p$  spaces. *Acta Mathematica*, 104 :93–140, 1960.

Lars Hörmander. On the uniqueness of the cauchy problem. *Mathematica Scandinavica*, 6 :213–225, 1958.

Lars Hörmander. Differential operators of principal type. *Mathematische Annalen*, 140 :124–146, 1960.

Lars Hörmander. *Linear Partial Differential Operators*. Springer Berlin, Heidelberg, 1963.

Lars Hörmander. *The analysis of linear partial differential Operators, I, II*. Springer Verlag, 1983.

Lars Hörmander. *The analysis of linear partial differential Operators, III, IV*. Springer Verlag, 1985.

L. Hörmander JJ. Duistermaat. *Fourier Integral Operators*. Birkhäuser Boston, MA, 1972.

Nicolas Lerner. A tribute to lars hörmänder. *Matapli*, 100, 2013.

# Introduction à l'analyse microlocale



Auriane GABAUT, Claire PLATEAUX, Enguérand TAMAGNA  
Encadré par Patrick GÉRARD  
Juin 2022

## Sommaire

---

|                                                              |           |
|--------------------------------------------------------------|-----------|
| <b>Introduction et Motivations</b>                           | <b>66</b> |
| <b>5.1 Rappels sur les distributions</b>                     | <b>66</b> |
| 5.1.1 Définitions                                            | 66        |
| 5.1.2 Convergence de distributions                           | 67        |
| 5.1.3 Dérivation de distributions                            | 67        |
| 5.1.4 Produit de distributions                               | 68        |
| 5.1.5 Transformation de Fourier                              | 68        |
| <b>5.2 Vers le Front d'Onde</b>                              | <b>69</b> |
| 5.2.1 Support et support singulier                           | 69        |
| 5.2.2 Le cône                                                | 71        |
| 5.2.3 Le front d'onde                                        | 72        |
| <b>5.3 Intégrales oscillantes</b>                            | <b>74</b> |
| 5.3.1 Quelques résultats utiles                              | 75        |
| 5.3.2 Construction de l'intégrale oscillante                 | 78        |
| 5.3.3 Front d'onde d'une intégrale oscillante                | 81        |
| <b>5.4 Du produit de distributions</b>                       | <b>83</b> |
| 5.4.1 Généralités                                            | 84        |
| 5.4.2 Front d'onde et produit                                | 86        |
| <b>5.5 Front d'onde et équations aux dérivées partielles</b> | <b>99</b> |
| 5.5.1 Théorème fondamental ( <i>Hörmander</i> )              | 100       |
| 5.5.2 Cas d'un opérateur elliptique                          | 103       |

---

---

## Introduction et Motivations

---

En 1983, le mathématicien suédois Lars Hörmander publie son premier volume de *The Analysis of Linear Partial Differential Operators* [1], dans lequel il introduit un nouvel objet mathématique très puissant pour décrire les singularités d'une distribution : le *front d'onde*.

Le principe est d'associer à chaque distribution sur un ouvert de  $\mathbb{R}^n$ , un sous-ensemble fermé de  $\Omega \times \mathbb{R}^n \setminus \{0\}$ , appelé *front d'onde* de la distribution, qui se projette sur le support singulier de la distribution, c'est-à-dire le complémentaire des points au voisinage desquels cette distribution est une fonction de classe  $C^\infty$ . Ce front d'onde possède non seulement de remarquables propriétés concernant les solutions d'équations aux dérivées partielles linéaires à coefficients  $C^\infty$ , mais il permet également de définir sur les distributions des opérations en général interdites, telles que leur produit.

Nous tenons à remercier chaleureusement M. Patrick Gérard de nous avoir encadré et de nous avoir transmis avec passion ces connaissances. Toute notre gratitude va à la direction des études de l'Université-Paris-Saclay et du Magistère de Mathématiques qui nous ont permis de compléter notre formation par un travail encadré de recherche.

*Remarque.* La rédaction du mémoire suit :

- GABAUT Auriane : 5.3
- PLATEAUX Claire : 5.1, 5.2 et 5.5
- TAMAGNA Enguérand : 5.4

---

### 5.1. Rappels sur les distributions

---

#### 5.1.1. Définitions

Étant donné un compact  $K$  de  $\mathbb{R}^n$ , nous noterons  $C_K^\infty$  l'espace des fonctions de classe  $C^\infty$  à support dans  $K$ . Si  $\Omega$  est un ouvert de  $\mathbb{R}^n$ , on note  $C_0^\infty(\Omega)$  l'espace des fonctions  $C^\infty$  dont le support est un compact de  $\Omega$ . C'est la réunion des  $C_K^\infty$  lorsque  $K$  parcourt l'ensemble des compacts inclus dans  $\Omega$ .

Si  $f$  est une fonction localement intégrable dans  $\Omega$  (c'est-à-dire intégrable sur tout compact de  $\Omega$ ), on peut lui faire correspondre la forme linéaire suivante sur  $C_0^\infty(\Omega)$  :

$$\forall \varphi \in C_0^\infty(\Omega), \langle f, \varphi \rangle = \int_{\Omega} f(x) \varphi(x) dx \quad (5.1)$$

et il est clair que l'on a la majoration :

$$\forall K \in \Omega, \forall \varphi \in C_K^\infty(\Omega), |\langle f, \varphi \rangle| \leq \left( \int_{\Omega} |f(x)| dx \right) \sup_{x \in K} |\varphi(x)| \quad (5.2)$$

#### Théorème 5.1.1 -

Deux fonctions localement intégrables dans  $\Omega$  définissent la même forme linéaire sur  $C_0^\infty(\Omega)$  si et seulement si elles sont égales presque partout.

#### Définition 5.1.1 -

L'espace  $\mathcal{D}'(\Omega)$  des distributions dans  $\Omega$  est constitué des formes linéaires  $u$  sur  $C_0^\infty(\Omega)$  vérifiant la propriété de continuité suivante :

Quel que soit  $K \in \Omega$ , il existe un entier  $p$  et une constante  $C$  tels que

$$\forall \varphi \in C_K^\infty(\Omega), |\langle u, \varphi \rangle| \leq C \sup_{x \in K, |\alpha| \leq p} |\partial^\alpha \varphi(x)| \quad (5.3)$$

Si l'on peut choisir l'entier  $p$  ci-dessus indépendamment de  $K$ , on dit que la distribution est d'ordre  $\leq p$  dans  $\Omega$ .

*Notation.* On note pour  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$  :

$$|\alpha| = \alpha_1 + \dots + \alpha_n \quad \text{et} \quad \partial^\alpha u = \frac{\partial^{|\alpha|} u}{\partial^{\alpha_1} x_1 \dots \partial^{\alpha_n} x_n}$$

On note aussi  $x^\alpha$  le monôme  $x_1^{\alpha_1} \dots x_n^{\alpha_n}$ .

### 5.1.2. Convergence de distributions

#### Définition 5.1.2 -

On dit qu'une suite  $(u_j)_{j \in \mathbb{N}}$  de distributions converge vers  $u$  si

$$\forall \varphi \in C_0^\infty(\Omega), \langle u_j, \varphi \rangle \xrightarrow{j \rightarrow \infty} \langle u, \varphi \rangle$$

*Rappel.* On rappelle qu'on peut démontrer que toute distribution est limite d'une suite de fonctions  $C^\infty$ .

### 5.1.3. Dérivation de distributions

#### Théorème 5.1.2 -

On généralise la formule d'intégration par parties aux distributions en posant pour  $u \in \mathcal{D}'(\Omega)$ , la distribution  $\partial u / \partial x_k$  définie par

$$\left\langle \frac{\partial u}{\partial x_k}, \varphi \right\rangle = - \left\langle u, \frac{\partial \varphi}{\partial x_k} \right\rangle.$$

La dérivation est continue : si  $u_j \rightarrow u$  dans  $\mathcal{D}'(\Omega)$ , alors  $\partial u_j / \partial x_k \rightarrow \partial u / \partial x_k$

*Démonstration.* Il faut juste vérifier que, la distribution  $u$  vérifiant la propriété de continuité (5.3), la forme linéaire  $\partial u / \partial x_k$  vérifie une propriété du même type. On a en effet, pour  $\phi \in C_k^\infty$ ,

$$\left| \left\langle \frac{\partial u}{\partial x_k}, \phi \right\rangle \right| = \left| \left\langle u, \frac{\partial \phi}{\partial x_k} \right\rangle \right| \leq C \sup_{x \in K, |\alpha| \leq p} \left| \partial^\alpha \frac{\partial \phi}{\partial x_k}(x) \right| \leq C \sup_{x \in K, |\beta| \leq p+1} |\partial^\beta \phi(x)|$$

Quand à la continuité, elle est évidente : on a

$$\left\langle \frac{\partial u_j}{\partial x_k}, \varphi \right\rangle = - \left\langle u_j, \frac{\partial \varphi}{\partial x_k} \right\rangle \rightarrow - \left\langle u, \frac{\partial \varphi}{\partial x_k} \right\rangle = \left\langle \frac{\partial u}{\partial x_k}, \varphi \right\rangle$$

ce qui exprime exactement que  $\partial u_j / \partial x_k \rightarrow \partial u / \partial x_k$  □

#### 5.1.4. Produit de distributions

*Définition 5.1.3 -*

Soient  $\Phi \in C^\infty(\Omega)$  et  $u \in \mathcal{D}'(\Omega)$ . La distribution  $\Phi u$  est définie par

$$\forall \varphi \in C_0^\infty(\Omega), \langle \Phi u, \varphi \rangle = \langle u, \Phi \varphi \rangle$$

La forme linéaire ainsi définie vérifie les bonnes estimations de continuité, et possède des propriétés agréables, comme la formule de Leibniz suivante.

*Proposition 5.1.2 - Formule de Leibniz*

Soit  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^d$  un multi-indice. Soit  $u \in \mathcal{D}'(\Omega)$  et  $\phi \in C^\infty(\Omega)$ . On a :

$$\partial^\alpha (u\phi) = \sum_{\beta \in \mathbb{N}^n, \beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta u \cdot \partial^{\alpha-\beta} \phi,$$

où  $\beta \leq \alpha$  signifie que  $\forall i = 1, \dots, n, \beta_i \leq \alpha_i$  et où  $\binom{\alpha}{\beta}$  désigne le coefficient multinomial :

$$\binom{\alpha}{\beta} = \binom{\alpha_1}{\beta_1} \dots \binom{\alpha_n}{\beta_n}.$$

#### 5.1.5. Transformation de Fourier

Cette partie vise à donner quelques rappels, sans démonstrations, qui seront utiles pour la suite.

*Définition 5.1.4 -*

On appelle  $\mathcal{S}'(\mathbb{R}^n)$  l'espace des distributions  $u \in \mathcal{D}'(\mathbb{R}^n)$  tels que, pour  $C > 0$  et  $N \in \mathbb{N}$  tel que

$$\forall \phi \in \mathcal{D}'(\mathbb{R}^n), |\langle u, \phi \rangle| \leq C \sup_{|\alpha|, |\beta| \leq N; x \in \mathbb{R}^n} |x^\alpha \partial^\beta \phi(x)|$$

- $\mathcal{S}'(\mathbb{R}^n)$  contient toutes les distributions à support compact qu'on note  $\mathcal{E}'(\mathbb{R}^n)$ .
- $\mathcal{S}'(\mathbb{R}^n)$  contient tous les espaces  $L^p$ .
- $\mathcal{S}'(\mathbb{R}^n)$  contient toutes les (classes de) fonctions localement intégrables qui sont majorées en module par un polynôme.
- $\mathcal{S}'(\mathbb{R}^n)$  est stable par dérivation, par multiplication par les polynômes.



Nous ne définirons pas, du moins dans le cas général, la transformée de Fourier  $u \mapsto \mathcal{F}u$ , souvent notée  $\hat{u}$ , mais nous en énoncerons les principales propriétés :

- $\mathcal{F}$  est un isomorphisme d'espaces vectoriels de  $\mathcal{S}'$  sur lui-même, on a  $\mathcal{F}^{-1} = \frac{1}{(2\pi)^d} \check{\mathcal{F}}$  où  $\check{\mathcal{F}} = \sigma \circ \mathcal{F}$  et

$$\sigma : \phi \in \mathcal{S}(\mathbb{R}^n) \mapsto (x \mapsto \phi(-x)) \in \mathcal{S}(\mathbb{R}^n)$$

- Pour  $u \in L^1(\mathbb{R}^n)$ , sa transformée de Fourier est la fonction continue bornée et tendant vers 0 à l'infini définie par la formule classique

$$\hat{u}(\xi) = \int e^{-ix \cdot \xi} u(x) dx.$$

- Si  $u$  est une distribution à support compact, sa transformée de Fourier est la fonction définie pour chaque  $\xi$  par

$$\hat{u}(\xi) = \langle u(x), e^{-ix \cdot \xi} \rangle \quad (5.4)$$

formule où  $x$  est une lettre muette et où on utilise l'extension, vue dans la section précédente, de  $\langle u, \cdot \rangle$  aux fonctions appartenant à  $C^\infty(\mathbb{R}^n)$ .

- La formule  $\mathcal{F}(\partial u / \partial x_j) = i\xi_j \hat{u}(\xi)$ , qui s'obtient habituellement en dérivant sous le signe somme, est valable en toute généralité.
- On a de même  $\mathcal{F}(x_j u) = i\partial \hat{u} / \partial \xi_j$  pour tout  $u \in \mathcal{S}'(\mathbb{R}^n)$ .

#### Théorème 5.1.3 -

Soit  $u$  une distribution à support compact.

- La fonction  $\hat{u}$  est de classe  $C^\infty$  et à croissance polynomiale ainsi que toutes ses dérivées : il existe des constantes  $C_\alpha$  et  $N_\alpha$  telles que l'on ait :

$$|\partial_\xi^\alpha \hat{u}(\xi)| \leq C_\alpha (1 + |\xi|)^{N_\alpha}$$

- Pour que  $u$  soit de classe  $C^\infty$ , il faut et il suffit que  $\hat{u}$  soit à décroissance rapide, i.e.

$$\forall N, \exists C, \forall \xi, \quad |\hat{u}(\xi)| \leq C(1 + |\xi|)^{-N}$$

## 5.2. Vers le Front d'Onse

### 5.2.1. Support et support singulier

#### Définition 5.2.1 -

Si  $\omega \subset \Omega$  sont deux ouverts, une fonction  $C^\infty$  à support dans un compact de  $\omega$  est a fortiori à support dans un compact de  $\Omega$ . Pour  $u \in D'(\Omega)$  sa restriction à  $\omega$  est la distribution  $u|_\omega$  définie par

$$\forall \varphi \in C_0^\infty(\omega), \langle u|_\omega, \varphi \rangle = \langle u, \varphi \rangle$$

Définition 5.2.2 -

Pour  $u \in \mathcal{D}'(\Omega)$  et  $x_0 \in \Omega$ , on dit que  $u$  est *nulle au voisinage de  $x_0$*  s'il existe un ouvert  $\omega \subset \Omega$  tel que  $u|_\omega = 0$ .

On appelle *support de  $u$*  l'ensemble fermé (relativement à  $\Omega$ ) noté  $\text{supp}(u)$  constitué des points  $x \in \Omega$  tels que  $u$  ne soit pas nulle au voisinage de  $x$ .

Théorème 5.2.1 -

La restriction de  $u$  à  $\Omega \setminus \text{supp}(u)$  est nulle.

Extension de la dualité :

Lorsque  $u \in \mathcal{D}'(\Omega)$  est à support compact, on peut définir  $\langle u, \phi \rangle$  pour  $\phi \in C^\infty(\Omega)$  de la façon suivante :

On choisit  $\psi \in \mathcal{D}'(\Omega)$  égale à 1 au voisinage de  $\text{supp}(u)$  et on pose  $\langle u, \phi \rangle = \langle u, \psi \phi \rangle$ , le membre de droite étant bien défini. Il faut bien sûr que le résultat ne change pas si l'on remplace  $\psi$  par une fonction analogue  $\psi'$ , ce qui évident puisque la fonction  $(\psi - \psi')\phi$  a son support inclus dans  $\Omega \setminus \text{supp}(u)$ .

Proposition 5.2.1 -

Soit  $u \in \mathcal{D}'(\Omega)$  à support compact, et soit  $K \subset \Omega$  un voisinage compact de  $\text{supp}(u)$ . Il existe alors  $p$  et  $C$  tels que l'on ait, pour l'extension ci-dessus,

$$\forall \varphi \in C^\infty(\Omega), |\langle u, \varphi \rangle| \leq C \sup_{x \in K, |\alpha| \leq p} |\partial^\alpha \varphi(x)|$$

*Démonstration.* On choisit  $\psi$  comme ci-dessus et à support dans  $K$ , on applique l'inégalité (5.3) à la fonction  $\psi \phi$  et on développe les  $\partial^\alpha(\psi \phi)$  à l'aide de la formule de Leibniz.  $\square$

*Rappel.* Toute distribution à support compact est d'ordre fini.

Définition 5.2.3 -

Pour  $u \in \mathcal{D}'(\Omega)$  et  $x_0 \in \Omega$ , on dit que  $u$  est  $C^\infty$  *au voisinage de  $x_0$*  s'il existe un ouvert  $\omega$  avec  $x_0 \in \omega \subset \Omega$  tel que  $u|_\omega \in C^\infty(\omega)$ .

On appelle *support singulier de  $u$*  l'ensemble fermé (relativement à  $\Omega$ ) noté  $\text{supp sing}(u)$  constitué des points  $x \in \Omega$  tels que  $u$  ne soit pas  $C^\infty$  au voisinage de  $x$ .

Théorème 5.2.2 -

En posant  $V = \Omega \setminus \text{supp sing}(u)$ , on a  $u|_V \in C^\infty(V)$ .

*Démonstration.* Par définition, il existe un recouvrement de  $V$  par des ouverts  $w_j$  tels que la restriction de  $u$  à  $w_j$  soit une fonction  $f_j \in C^\infty(w_j)$ . Pour  $\phi \in \mathcal{D}'(w_j \cap w_k)$  on a donc  $\langle u, \phi \rangle = \int f_j \phi dx = \int f_k \phi dx$  ce qui entraîne que  $f_j$  et  $f_k$  coïncident dans  $w_j \cap w_k$ . On peut donc recoller ces fonctions et définir  $f \in C^\infty(V)$  en décidant que  $f(x)$  est la valeur commune de tous les  $f_j(x)$  pour  $w_j \ni x$ .  $\square$

Le théorème (3) permet de déduire :

Corollaire 5.2.2 -

Soit  $u \in \mathcal{D}'(\Omega)$  et  $x_0 \in \Omega$ . Les quatres propriétés suivantes sont équivalentes :

- (i)  $x_0 \notin \text{supp sing}(u)$ .
- (ii)  $u$  est  $C^\infty$  au voisinage de  $x_0$ .
- (iii) Il existe  $\phi \in \mathcal{D}'(\Omega)$  telle que  $\phi(x_0) \neq 0$  et que  $\widehat{\phi u}(\xi)$  soit à décroissance rapide.
- (iv) Il existe un ouvert  $w$  avec  $x_0 \in w \subset \Omega$  tel que, pour tout  $\phi \in \mathcal{D}'(\Omega)$ , la fonction  $\widehat{\phi u}(\xi)$  soit à décroissance rapide.

## 5.2.2. Le cône

Définition 5.2.4 -

On définit, pour  $v \in \mathcal{E}'(\Omega)$ , le cône  $\Sigma(v)$  comme l'ensemble des  $\eta \in \mathbb{R}^n \setminus \{0\}$  au voisinage desquels  $\hat{v}$  n'est pas à décroissance rapide. C'est un fermé de  $\mathbb{R}^n \setminus \{0\}$ , et on a  $\Sigma(v) = \emptyset$  si et seulement si  $v \in C_0^\infty(\Omega)$ .

Lemme 5.2.2 -

Pour  $\phi \in C_0^\infty(\mathbb{R}^n)$  et  $v \in \mathcal{E}'(\mathbb{R}^n)$ , on a

$$\Sigma(\phi v) \subset \Sigma(v) \quad (5.5)$$

*Démonstration.* La transformée de Fourier de  $u = \phi v$  est la convolution

$$\hat{u}(\xi) = (2\pi)^{-n} \int \hat{\phi}(\eta) \hat{v}(\xi - \eta) d\eta$$

où  $\hat{\phi} \in \mathcal{S}$ . Pour un certain  $M \geq 0$  on a

$$|\hat{v}(\xi)| \leq C(1 + |\xi|)^M.$$

Soit  $c \in ]0, 1[$ . On va couper l'intégrale en deux en fonction de si  $|\eta| < c|\xi|$  et  $|\eta| \geq c|\xi|$ . On a alors :

$$\begin{aligned} (2\pi)^n |\hat{u}(\xi)| &\leq \int_{|\eta| < c|\xi|} |\hat{\phi}(\eta) \hat{v}(\xi - \eta)| d\eta + \int_{|\eta| \geq c|\xi|} |\hat{\phi}(\eta) \hat{v}(\xi - \eta)| d\eta \\ &\leq \sup_{|\xi - \eta| < c|\xi|} |\hat{v}| \|\hat{\phi}\|_{L^1} + \int_{|\eta| \geq c|\xi|} |\hat{\phi}(\eta) \hat{v}(\xi - \eta)| d\eta \quad (\text{changement de variable}) \\ &\leq \sup_{|\xi - \eta| < c|\xi|} |\hat{v}| \|\hat{\phi}\|_{L^1} + C \int_{|\eta| \geq c|\xi|} |\hat{\phi}(\eta)| (1 + |\eta|)^M (1 + c^{-1})^M d\eta \end{aligned} \quad (5.6)$$

car si  $|\eta| \geq c|\xi|$ , alors  $|\xi - \eta| \leq (1 + c^{-1})|\eta|$  et donc :

$$\begin{aligned} |\hat{v}(\xi - \eta)| &\leq C(1 + |\xi - \eta|)^M \\ &\leq C(1 + (1 + c^{-1})|\eta|)^M \\ &\leq C(1 + |\eta|)^M \left(1 + \frac{|\eta|c^{-1}}{1 + |\eta|}\right)^M \\ &\leq C(1 + |\eta|)^M (1 + c^{-1})^M \end{aligned}$$

Si  $\Gamma$  est un cône ouvert dans lequel  $\hat{v}$  est à décroissance rapide, et  $\Gamma_1$  est un cône fermé de  $\Gamma \cup \{0\}$  on peut choisir  $c$  tel que  $\eta \in \Gamma$  si  $\xi \in \Gamma_1$  et  $|\xi - \eta| < c|\xi|$ , ce qui est évidemment possible si  $|\xi| = 1$ . Puisque

$|\eta| \geq (1-c)|\xi|$ , il suit de (5.6) que  $\hat{u}$  est rapidement décroissante dans  $\Gamma_1$ . En fait, on a pour  $N \geq 0$  :

$$\begin{aligned} \sup_{\Gamma_1} (1+|\xi|)^N |\hat{u}(\xi)| &\leq (1-c)^{-N} \sup_{\Gamma} |\hat{v}(\eta)| (1+|\eta|)^N \|\hat{\phi}\|_{L^1} \\ &\quad + C(1+c^{-1})^{N+M} \int |\hat{\phi}(\eta)| (1+|\eta|)^{N+M} d\eta \end{aligned}$$

ce qui conclue la preuve car cette quantité est finie.  $\square$

Définition 5.2.5 -

Si  $\Omega$  est un ouvert de  $\mathbb{R}^n$  et  $u \in \mathcal{D}'(\Omega)$ , on pose pour  $u \in \Omega$  :

$$\Sigma_x(u) = \bigcap_{\phi} \Sigma(\phi u); \quad \phi \in \mathcal{D}(\Omega), \phi(x) \neq 0$$

Lemme 5.2.2 -

On a  $\Sigma(\phi u) \rightarrow \Sigma_x(u)$  si  $\phi \in \mathcal{D}(\Omega), \phi(x) \neq 0$  et  $\text{supp}(\phi) \rightarrow \{x\}$

*Démonstration.* En effet, si  $V$  est un cône ouvert  $\supset \Sigma_x(u)$ , la compacité de la sphère unité permet de trouver  $\phi_1(x) \dots \phi_j(x) \neq 0$  tels que  $\bigcap_1^j \Sigma(\phi_i u) \subset V$ .

Soit  $\phi \in \mathcal{D}(\Omega)$  tel que  $\text{supp}(\phi)$  est assez proche de  $x$  pour que  $\phi_1 \dots \phi_j \neq 0$ , alors on peut écrire  $\phi = \psi \phi_1 \dots \phi_j$  avec  $\psi \in \mathcal{D}(\Omega)$  et obtenir d'après (5.5) que

$$\Sigma(\phi u) \subset \bigcap_1^j \Sigma(\phi_i u) \subset V.$$

Enfin par définition,  $\Sigma(\phi u) \supset \Sigma_x(u)$  lorsque  $\phi(x) \neq 0$  ce qui conclue la preuve.  $\square$

*Remarque.* En particulier, on a que  $\Sigma_x(u) = \emptyset$  si et seulement si  $\phi u \in C^\infty$  pour une certaine fonction  $\phi \in \mathcal{D}(\Omega)$  tel que  $\phi(x) \neq 0$  i.e.  $x \notin \text{supp sing}(u)$ .

### 5.2.3. Le front d'onde

Le front d'onde de  $u \in \mathcal{D}'(\Omega)$  va être un sous-ensemble conique de  $\Omega \times (\mathbb{R}^n \setminus \{0\})$  qui va décrire non seulement les points où  $u$  est singulière, mais encore les directions dans lesquelles elle est singulière.

Définition 5.2.6 -

Nous dirons qu'un sous-ensemble  $A \subset \mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\})$  est *conique* (ou *positivement homogène en  $\xi$* ) si, pour  $t > 0$ , on a

$$(x, \xi) \in A \iff (x, t\xi) \in A.$$

Définition 5.2.7 -

Soit  $u \in \mathcal{D}'(\Omega)$ . On dit que  $u$  est *microlocalement de classe  $C^\infty$*  en un point  $(x_0, \xi_0) \in \Omega \times (\mathbb{R}^n \setminus \{0\})$  s'il existe un ouvert  $\omega$  vérifiant  $x_0 \in \omega \subset \Omega$  et un cône ouvert  $\Gamma$  de  $\mathbb{R}^n \setminus \{0\}$  contenant  $\xi_0$  tels que l'on ait

$$\forall \varphi \in C_0^\infty(\omega), \forall N, \exists C |\widehat{\varphi u}(\xi)| \leq C(1+|\xi|)^{-N} \text{ pour } \xi \in \Gamma \quad (5.7)$$

L'ensemble des  $(x_0, \xi_0)$  où  $u$  n'est pas microlocalement de classe  $C^\infty$  est un fermé conique de  $\Omega \times (\mathbb{R}^n \setminus \{0\})$  appelé le *front d'onde* de  $u$  et noté  $WF(u)$ .

**Exemple 5.2.1.** Dans  $\mathbb{R}^n$ , le front d'onde de  $\delta_0$  est  $0 \times (\mathbb{R}^n \setminus \{0\})$ .

▷ En effet, si  $x_0 \neq 0$ , on peut trouver un ouvert  $w$  contenant  $x_0$  mais pas 0 et alors les produits  $\widehat{\phi\delta}$  seront nuls. Ainsi,  $\delta_0$  est microlocalement  $C^\infty$  en  $x_0$ .

Par l'absurde, si pour  $\xi \in \mathbb{R}^n$ ,  $(0, \xi) \notin WF(\delta_0)$ , alors  $\exists w \subset \Omega$  contenant 0 tel que  $\forall \phi \in C_0^\infty(w)$ ,  $\forall N$ ,  $\exists C$ ,

$$|\widehat{\phi\delta_0}(\xi)| \leq C(1 + |\xi|)^{-N} \quad \forall \xi \in \Gamma$$

Si on choisit  $\phi$  tel que  $\phi(0) \neq 0$ , alors  $\widehat{\phi\delta_0} = \phi(0)$  ne décroît dans aucune direction ce qui est absurde.  $\square$

Théorème 5.2.3 -

| La projection sur  $\mathbb{R}^n$  de  $WF(u)$  est  $supp \, sing(u)$ .

*Démonstration.* Soit  $x_0 \in \Omega$  tel que  $x_0 \notin supp \, sing(u)$ . Alors il existe un ouvert  $w$  avec  $x_0 \in w \subset \Omega$  tel que  $\forall \phi \in \mathcal{D}'(\Omega)$ ,  $\widehat{\phi u}(\xi)$  soit à croissance rapide i.e.

$$\forall N, \exists C, |\widehat{\phi u}(\xi)| \leq C(1 + |\xi|)^{-N} \quad \forall \xi \in \mathbb{R}^n.$$

Donc  $u$  est microlocalement  $C^\infty$  en  $(x_0, \xi_0) \forall \xi_0 \in \mathbb{R}^n \setminus \{0\}$  ce qui signifie que  $(x_0, \xi_0) \notin WF(u)$ .

Réciproquement, si  $(x_0, \xi_0) \notin WF(u) \forall x_0 \in \Omega, \xi_0 \in \mathbb{R}^n \setminus \{0\}$ , alors  $u$  est microlocalement de classe  $C^\infty$  en ce point. Il existe donc  $w$  ouvert et un cône  $\Gamma \ni \xi_0$  tel que

$$\forall \phi \in \mathcal{D}'(\Omega), \forall N, \exists C, |\widehat{\phi u}(\xi)| \leq C(1 + |\xi|)^{-N} \quad \forall \xi \in \Gamma.$$

Le théorème de Borel-Lebesgue appliqué à la sphère de  $\mathbb{R}^n$  permet de trouver un nombre fini de tels couples  $(w_j, \Gamma_j)$  de manière que les  $\Gamma_j$  recouvrent  $\mathbb{R}^n \setminus \{0\}$ . Pour  $\phi \in \mathcal{D}'(\cap_j w_j)$  la fonction  $\widehat{\phi u}$  est à décroissance rapide, donc  $x_0 \notin supp \, sing(u)$  ce qui achève la démonstration.  $\square$

Proposition 5.2.3 -

| Si  $u \in \mathcal{E}'(\mathbb{R}^n)$ , alors la projection de  $WF(u)$  sur la seconde variable notée  $W$  est  $\Sigma(u)$ .

*Démonstration.* La projection  $W$  est contenue dans  $\Sigma(u)$  par définition du front d'onde. De plus, elle est fermée puisque son intersection avec la sphère unité est la projection d'un compact dans  $\mathbb{R}^n \times S^{n-1}$ . Si  $V$  est un voisinage conique de  $W$  alors tout  $x \in \mathbb{R}^n$  à un voisinage  $U_x$  tel que

$$\Sigma(\phi u) \subset V \quad \text{si } \phi \in \mathcal{D}(U_x).$$

On peut recouvrir  $supp \, (u)$  par un nombre fini de tels voisinages  $U_x$ , et choisir  $\phi_j \in \mathcal{D}(U_{x_j})$  tel que  $\Sigma\phi_j = 1$  près de  $supp \, (u)$ . Mais alors il suit que

$$\Sigma(u) = \Sigma(\sum \phi_j u) \subset \bigcup \Sigma(\phi_j u) \subset V$$

ce qui prouve la proposition.  $\square$

Proposition 5.2.3 -

| Pour tout  $f \in C_0^\infty(\Omega)$ , on a  $WF(fu) \subset WF(u)$ .

*Démonstration.* Il suffit de montrer que pour tout  $x \in \Omega$ ,  $\Sigma_x(fu) \subset \Sigma_x(u)$ .

Soit  $x \in \Omega$ . Prenons  $\chi \in \mathbb{C}_0^\infty$  égale à 1 près de  $x$  et  $\chi_1$  égale à 1 au voisinage de  $\text{supp } (\chi)$ . On a alors d'après (5.5) que

$$\Sigma_x(fu) \subset \Sigma(\chi fu) = \Sigma(\chi f \chi_1 u) \subset \Sigma(f \chi_1 u) \subset \Sigma(\chi_1 u)$$

Le résultat suit en faisant tendre  $\text{supp } \chi_1$  vers  $\{x\}$ .  $\square$

*Proposition 5.2.3 -*

On a, pour tout  $\alpha$ ,  $\text{WF}(\partial^\alpha u) \subset \text{WF}(u)$  et  $\text{WF}(Pu) \subset \text{WF}(u)$ , pour  $P$  opérateur différentiel linéaire à coefficients  $C^\infty$ .

*Démonstration.* Soit  $x \in \Omega$ . Prenons  $\chi \in \mathbb{C}_0^\infty$  égale à 1 près de  $x$  et  $\chi_1$  égale à 1 au voisinage de  $\text{supp}(\chi)$ . On a alors d'après (5.5) que

$$\Sigma_x(\partial^\alpha u) \subset \Sigma(\chi \partial^\alpha u) = \Sigma(\chi \partial^\alpha \chi_1 u) \subset \Sigma(\partial^\alpha \chi_1 u) \subset \Sigma(\chi_1 u)$$

La dernière inclusion vient du fait que si  $\widehat{\chi_1 u}$  est à décroissance rapide, alors  $\widehat{\partial^\alpha \chi_1 u} = (i)^{|\alpha|} \xi^\alpha \widehat{\chi_1 u}$  aussi. Le résultat suit en faisant tendre  $\text{supp } \chi_1$  vers  $\{x\}$ .

En sommant les  $\partial^\alpha$  et en utilisant la proposition précédente, on obtient le résultat pour tout  $P$  opérateur différentiel linéaire à coefficients  $C^\infty$ .  $\square$

*Remarque.* (2.3) Pour vérifier que  $(x_0, \xi_0) \notin \text{WF}(u)$ , il suffit de vérifier que pour une seule fonction  $\psi \in \mathcal{D}(\Omega)$  telle que  $\psi(x_0) \neq 0$ , la transformée de Fourier de  $\psi u$  est à décroissance rapide dans un cône contenant  $\xi_0$ . En effet, si  $w$  est un ouvert contenant  $x_0$  dans lequel  $\psi$  ne s'annule pas, tout  $\phi \in \mathcal{D}(\Omega)$  pourra s'écrire  $\theta\psi$  avec  $\theta \in \mathcal{D}(\Omega)$ . Nous venons de voir que la décroissance rapide de  $\widehat{\psi u}$  dans un voisinage conique de  $\xi_0$  implique celle de  $\widehat{\theta\psi u}$ .

---

## 5.3. Intégrales oscillantes

---

L'objet de cette partie est d'étudier les intégrales du type :

$$\int e^{i\phi(x, \theta)} a(x, \theta) d\theta$$

appelées intégrales oscillantes et définies pour des fonctions  $\phi$ , et  $a$  assez régulières. Nous voulons déterminer les conditions sur  $a$  pour que l'intégrale ci-dessus définisse une distribution dont nous pourrions alors discuter du front d'onde.

## 5.3.1. Quelques résultats utiles

Définition 5.3.1 -

Soient  $m \in \mathbb{R}$ , et soit  $X \subset \mathbb{R}^n$  un ouvert et  $\Gamma$  un cône ouvert de  $X \times (\mathbb{R}^N \setminus \{0\})$  pour  $N \in \mathbb{N}^*$ . On note  $S_{1,0}^m(X \times \mathbb{R}^N)$  l'ensemble des fonctions  $a \in C^\infty(X \times \mathbb{R}^N)$  telles que pour tout compact  $K \subset X$  et tout  $\alpha \in \mathbb{N}^N$ ,  $\beta \in \mathbb{N}^n$  :

$$\exists C_{\alpha,\beta,K} > 0, \quad |D_x^\beta D_\theta^\alpha a(x, \theta)| \leq C_{\alpha,\beta,K} (1 + |\theta|)^{m-|\alpha|}, \quad \forall x \in K, \theta \in \mathbb{R}^N$$

*Remarque.* Muni de ces semi-normes, l'espace  $S_{0,1}^m$  est un espace de Fréchet, i.e. qu'il s'agit d'un espace muni d'une famille de semi-normes séparantes, rendant l'espace complet.

Nous allons commencer par démontrer le lemme de la phase non stationnaire, nécessaire à la construction des intégrales oscillantes.

Lemme 5.3.0 -

Soit  $K \subset \mathbb{R}^n$  un compact, et  $X$  un voisinage ouvert de  $K$ . Soit  $k \geq 0$ .  
Si  $u \in C_0^k(K)$ ,  $f \in C^{k+1}(K)$  à valeurs réelles sur  $X$  tel que  $|\nabla f| \neq 0$  sur  $K$ . Alors, il existe une constante  $C$ , bornée lorsque  $f$  reste bornée dans  $C^{k+1}(X)$  tel que :

$$\forall \omega > 0, \quad \omega^k \left| \int u(x) e^{i\omega f(x)} dx \right| \leq C \sum_{|\alpha| \leq k} \sup |D^\alpha u| |f'|^{|\alpha|-2k} \quad (5.8)$$

*Démonstration.* Soit l'opérateur différentiel  $L = \frac{1}{|\nabla f|^2} \sum_{j=1}^N \partial_j f \partial_j \cdot$  de sorte que :

$$e^{i\omega f} = \frac{1}{i\omega} L(e^{i\omega f})$$

On peut donc écrire l'intégration par partie :

$$\begin{aligned} \int u(x) e^{i\omega f(x)} dx &= \frac{1}{i\omega} \int u(x) L(e^{i\omega f(x)}) dx \\ &= -\frac{1}{i\omega} \int e^{i\omega f(x)} ({}^t L)(u) dx \end{aligned}$$

où  ${}^t L \cdot = - \sum_{j=1}^n \partial_j \left( \frac{\partial_j f}{|\nabla f|^2} \times \cdot \right)$ . On obtient, en itérant l'intégration par partie  $k$  fois, l'égalité :

$$\begin{aligned} \omega^k \left| \int u e^{i\omega f} \right| &= \omega^{k-1} \left| \int ({}^t L)(u) e^{i\omega f} \right| \\ &= \left| \int ({}^t L)^k(u) e^{i\omega f} \right| \end{aligned}$$

Montrons maintenant par récurrence sur  $k \in \mathbb{N}^*$  que : "pour  $f \in C^{k+1}(K)$ ,  $u \in C_0^k(K)$ ,  $({}^t L)^k$  peut s'écrire comme  $\sum_{|\alpha| \leq k} a_{\alpha,k} \partial^\alpha$  où  $a_{\alpha,k}$  est une somme finie de polynômes en les dérivées de  $f$  d'ordre  $1 \leq \gamma \leq k+1$  multiplié par un facteur homogène de degré  $|\alpha| - 2k$  en  $|\nabla f|$ "

D'abord, si  $k = 1$ , on a :

$${}^t L = - \sum_{j=1}^n \underbrace{\frac{\partial_j f}{|\nabla f|^2}}_{=a_{e_j,1}} \partial_j \cdot + \underbrace{\frac{\partial_j^2 f |\nabla f|^2 - \partial_j f \partial_j |\nabla f|^2}{|\nabla f|^4}}_{=a_{0,1}} \times \cdot$$

On remarque alors que :

- $a_{e_j,1} = \partial_j f \times \frac{1}{|\nabla f|^2}$  où  $\partial_j f$  est bien un polynôme en les dérivées de  $f$  d'ordre 1 ou 2, et la fraction rationnelle  $\frac{1}{|\nabla f|^2}$  est homogène en  $|\nabla f|$  de degrés  $-1 = |\alpha| - 2k$  ;
- $a_{0,1} = \underbrace{\sum_{j=1}^n \left( \partial_j^2 f |\nabla f|^2 - \partial_j f \partial_j |\nabla f|^2 \right)}_{=p} \times \frac{1}{|\nabla f|^4}$  où  $p$  est un polynôme en les dérivées de  $f$  d'ordre compris entre 1 et 2 et  $\frac{1}{|\nabla f|^4}$  est homogène de degré  $-2 = |\alpha| - 2k$ .

On a donc bien la propriété vérifiée lorsque  $k = 1$ , maintenant supposons la propriété vraie jusqu'à un certain rang  $k \in \mathbb{N}^*$ . On peut alors écrire :

$$({}^t L)^k = \sum_{|\alpha| \leq k} a_{\alpha,k} \partial^\alpha.$$

avec  $a_{\alpha,k} = \sum_{i \in I_{\alpha,k}} P_{\alpha,k}^i \times H_{\alpha,k}^i$  où  $I_{\alpha,k}$  est fini,  $P_{\alpha,k}^i$  est un polynôme en les dérivées de  $f$  d'ordre compris entre 1 et  $k+1$  et  $H_{\alpha,k}^i$  est homogène de degré  $|\alpha| - 2k$  en  $|\nabla f|$  et :

$$\begin{aligned} ({}^t L)^{k+1} &= ({}^t L) \circ ({}^t L)^k \\ &= \left( - \sum_{j=1}^n \frac{\partial_j f}{|\nabla f|^2} \partial_j \cdot + a_{0,1} \times \cdot \right) \circ \left( \sum_{|\alpha| \leq k} a_{\alpha,k} \partial^\alpha \right) \\ &= \sum_{|\alpha| \leq k} \sum_{j=1}^n - \frac{\partial_j f}{|\nabla f|^2} \partial_j (a_{\alpha,k} \partial^\alpha) + a_{0,1} a_{\alpha,k} \partial^\alpha \\ &= \sum_{|\alpha| \leq k} \sum_{j=1}^n - \frac{\partial_j f}{|\nabla f|^2} \partial_j a_{\alpha,k} \partial^\alpha - \frac{\partial_j f}{|\nabla f|^2} a_{\alpha,k} \partial_j \partial^\alpha + a_{0,1} a_{\alpha,k} \partial^\alpha \end{aligned}$$

Si on note  $({}^t L)^{k+1} = \sum_{|\beta| \leq k+1} a_{\beta,k+1} \partial^\beta$  on peut alors distinguer trois cas :

- $|\beta| = k+1$ , alors il existe un  $\alpha$  telle que  $|\alpha| = k$  et  $\beta = \alpha + e_j$  pour un certain  $j \in \{1, \dots, n\}$ , et donc, par identification :

$$\begin{aligned} a_{\beta,k+1} &= - \frac{\partial_j f}{|\nabla f|^2} a_{\alpha,k} \\ &= \sum_{i \in I_{\alpha,k}} \underbrace{- \partial_j f \times P_{\alpha,k}^i}_{=P_{\beta,k+1}^i} \times \underbrace{\frac{H_{\alpha,k}^i}{|\nabla f|^2}}_{=H_{\beta,k+1}^i} \end{aligned}$$

donc  $P_{\beta,k+1}^i$  est bien un polynôme en les dérivées de  $f$  d'ordre compris entre 1 et  $k+1 \leq k+2$ , et  $H_{\beta,k+1}^i$  est homogène de degrés  $|\alpha| - 2k - 1 = |\beta| - 1 - 2k - 1 = |\beta| - 2(k+1)$  en  $|\nabla f|$  ;

- $|\beta| = 0$ , alors on retrouve les dérivées d'ordre 0 dans les termes de la somme seulement, de sorte que :

$$\begin{aligned} a_{\beta,k+1} &= - \sum_{j=1}^n \frac{\partial_j f}{|\nabla f|^2} \partial_j a_{\alpha,k} + a_{0,1} a_{\alpha,k} \\ &= \sum_{i \in I_{\alpha,k}} \underbrace{- \sum_{j=1}^n \frac{\partial_j f}{|\nabla f|^2} \partial_j P_{\alpha,k}^i H_{\alpha,k}^i}_{=A^i} - \sum_{j=1}^n \underbrace{\frac{\partial_j f}{|\nabla f|^2} P_{\alpha,k}^i \partial_j H_{\alpha,k}^i}_{=B^i} + \sum_{l \in I_{0,k}} \underbrace{P_{0,1}^l H_{0,1}^l P_{\alpha,k}^i H_{\alpha,k}^i}_{=C^i} \end{aligned}$$



- (a)  $A^i = - \sum_{j=1}^n \partial_j f \partial_j |\nabla f|^2 P_{\alpha,k}^i \times \frac{H_{\alpha,k}^i}{|\nabla f|^4}$  est bien une somme finie de polynômes en les dérivées de  $f$  d'ordre compris entre 1 et  $k+2$  multiplié par un facteur homogène de degré au moins  $|\alpha| - 2k - 2 = -2k - 1 = |\beta| - 2(k+1)$
- (b)  $B^i = - \sum_{j=1}^n \partial_j f P_{\alpha,k}^i \times \frac{\partial_j H_{\alpha,k}^i}{|\nabla f|^2}$  est bien une somme finie de polynômes en les dérivées de  $f$  d'ordre compris entre 1 et  $k+1 \leq k+2$  multiplié par un facteur homogène de degré au moins  $-1 + |\alpha| - 2k - 1 = |\beta| - 2(k+1)$
- (c)  $C^i = \sum_{l \in I_{0,1}} P_{0,1}^l P_{\alpha,k}^i |\nabla f|^2 \times \frac{H_{0,1}^l H_{\alpha,k}^i}{|\nabla f|^2}$  est bien une somme finie de polynômes en les dérivées de  $f$  d'ordre compris entre 1 et  $k+1 \leq k+2$  multiplié par un facteur homogène de degré au moins  $-1 + |\alpha| - 2k - 1 = |\beta| - 2(k+1)$
- $\frac{1}{|\beta|} \leq |\alpha| \leq k$ , alors il existe  $\alpha_1, \alpha_2$  tel que  $|\alpha_1| = |\beta|$ , et  $\alpha_2 + e_j = \beta$  pour un certain  $j \in \{1, \dots, n\}$ , de sorte que par identification :

$$\begin{aligned} a_{\beta,k+1} &= \sum_{j=1}^n -\frac{\partial_j f}{|\nabla f|^2} \partial_j a_{\alpha_1,k} - \frac{\partial_j f}{|\nabla f|^2} a_{\alpha_2,k} + a_{0,1} a_{\alpha_1,k} \\ &= \sum_{i \in I_{\alpha_1,k}} \sum_{j=1}^n \partial_j f \partial_j P_{\alpha_1,k}^i |\nabla f|^2 \frac{H_{\alpha_1,k}^i}{|\nabla f|^4} - P_{\alpha_1,k}^i \partial_j f \frac{\partial_j H_{\alpha_1,k}^i}{|\nabla f|^2} + \sum_{l \in I_{0,1}} |\nabla f|^2 P_{0,1}^l P_{\alpha_1,k}^i H_{0,1}^l \frac{H_{\alpha_1,k}^i}{|\nabla f|^2} \\ &\quad + \sum_{i \in I_{\alpha_2,k}} \sum_{j=1}^n -\partial_j f P_{\alpha_2,k}^i \frac{H_{\alpha_2,k}^i}{|\nabla f|^2} \end{aligned}$$

et de même que pour le cas précédemment, on remarque que  $a_{\beta,k+1}$  s'écrit comme une somme finie de polynômes en les dérivées d'ordre au plus  $k+2$  et de facteur homogène en  $|\nabla f|$  de degré  $|\alpha_1| - 2k - 2 = |\beta| - 2(k+1)$ , et  $|\alpha_2| - 2k - 1 = |\beta| - 2(k+1)$

Donc, dans tout les cas, le résultat est de nouveau vérifié au rang  $k+1$ , ce qui conclue la récurrence.

On peut donc considérer :

$$({}^t L)^k = \sum_{|\alpha| \leq k} a_{\alpha,k} \partial^\alpha$$

avec  $a_{\alpha,k} = \sum_{i \in I_{\alpha,k}} P_{\alpha,k}^i \times H_{\alpha,k}^i$  où  $I_{\alpha,k}$  est fini,  $P_{\alpha,k}^i$  est un polynôme en les dérivées de  $f$  d'ordre compris entre 1 et  $k+1$  et  $H_{\alpha,k}^i$  est homogène de degré  $|\alpha| - 2k$  en  $|\nabla f|$ ; on a alors :

$$\begin{aligned} \omega^k \left| \int u e^{i\omega f} \right| &= \left| \int ({}^t L)^k(u) e^{i\omega f} \right| \\ &\leq \int_K \sum_{|\alpha| \leq k} |a_{\alpha,k} \partial^\alpha u| \\ &= \int_K \sum_{|\alpha| \leq k} \sum_{i \in I_{\alpha,k}} \left| P_{\alpha,k}^i H_{\alpha,k}^i \partial^\alpha u \right| \end{aligned}$$

et  $f \in C^{k+1}(K)$ , donc par continuité sur le compact  $K$ , on peut majorer tout les  $P_{\alpha,k}^i$  par une constante  $C_{\alpha,k,f}$  qui reste bornée lorsque  $f$  l'est dans  $C^{k+1}(X)$ ; ensuite,  $H_{\alpha,k}^i$  est homogène de degré  $|\alpha| - 2k$  en  $|\nabla f|$  donc on peut les majorer par  $C'_{\alpha,k,f} |\nabla f|^{|\alpha| - 2k}$  de sorte que  $C'_{\alpha,k,f}$  reste bornée lorsque  $f$  le reste dans  $C^{k+1}(X)$  par continuité sur le compact  $K$ . On a donc finalement :

$$\begin{aligned} \omega^k \left| \int u e^{i\omega f} \right| &\leq |K| \sum_{|\alpha| \leq k} C_{\alpha,k,f} C'_{\alpha,k,f} \times \sup |\partial^\alpha u| |\nabla f|^{|\alpha| - 2k} \\ &\leq C_{K,k,f} \sup |D^\alpha u| |\nabla f|^{|\alpha| - 2k} \end{aligned}$$

Ce qui conclue la preuve du lemme de phase non stationnaire.  $\square$

### 5.3.2. Construction de l'intégrale oscillante

Nous pouvons à présent construire l'intégrale oscillante, et justifier dans quel cas il s'agit bien d'une distribution.

Définition 5.3.2 -

Soit  $X \subset \mathbb{R}^n$  un ouvert, et  $\Gamma$  un cône ouvert de  $X \times (\mathbb{R}^N \setminus \{0\})$ . On appelle fonction de phase sur  $\Gamma$  une fonction  $\phi \in C^\infty(\Gamma)$  telle que :

1.  $\phi(x, t\theta) = t\phi(x, \theta)$ ,  $\forall (x, \theta) \in \Gamma, t > 0$
2.  $d\phi \neq 0$  sur  $\Gamma$

Théorème 5.3.1 -

Soit  $\phi$  une fonction de phase sur  $\Gamma \subset X \times \mathbb{R}^N$ , et soit  $F$  un cône fermé sur  $\Gamma \cup (X \times \{0\})$ . Soit  $m \in \mathbb{R}$ , alors on peut définir de manière unique

$$I_\phi(a) : u \in C_0^\infty(X) \mapsto \int e^{i\phi(x, \theta)} a(x, \theta) u(x) dx d\theta$$

à tout  $a \in S_{0,1}^m$  et  $u \in C_0^\infty(X)$  de sorte que :

- i  $u \mapsto I_\phi(au)$  est une distribution d'ordre  $\leq k$  si  $a \in S_{0,1}^m$  et  $k > N + m$
- ii  $a \in S_{0,1}^m \mapsto I_\phi(au) \in \mathcal{D}'(X)$  est une application linéaire continue pour tout  $u \in C_0^\infty(X)$  fixé

*Démonstration.* Prenons dans un premier temps  $\chi \in C_0^\infty(\mathbb{R}^N)$  telle que  $\chi \equiv 1$  sur  $B_N(0, 1)$  et  $\chi \equiv 0$  en dehors de  $B_N(0, 2)$ . On pose :

$$\begin{cases} \chi_\nu(\theta) &= \chi(2^{-\nu}\theta) - \chi(2^{1-\nu}\theta) \\ \chi_0(\theta) &= \chi(\theta) \end{cases}, \quad \forall \theta \in \mathbb{R}^N, \nu \in \mathbb{N}^*$$

On remarque alors que pour  $\nu > 1$  :

$$\chi_\nu(\theta) = \chi_{\nu-1}(2^{-1}\theta) = \chi_1(2^{-\nu+1}\theta), \quad \forall \theta \in \mathbb{R}^N$$

dont on déduit que

$$\text{supp} \chi_\nu \subset [-2 \times 2^\nu; 2 \times 2^\nu] \quad \text{et} \quad \chi_\nu \equiv 0 \text{ sur } [-2^{\nu-1}; 2^{\nu-1}]$$

du fait que, comme  $\chi_1(\theta) = \chi\left(\frac{\theta}{2}\right) - \chi(\theta)$ ,  $\forall \theta \in \mathbb{R}^N$ , on a :

$$\text{supp} \chi_1 \subset \text{supp} \chi \left( \frac{\cdot}{2} \right) \subset [-2 \times 2; 2 \times 2] \quad \text{et} \quad \left( \theta \in [-1; 1] \Rightarrow \chi\left(\frac{\theta}{2}\right) = \chi(\theta) = 1 \Rightarrow \chi_1(\theta) = 0 \right)$$

Prenons  $N \in \mathbb{N}^*$  et  $\theta \in \mathbb{R}^N$  :

$$\begin{aligned} \sum_{\nu=0}^N \chi_\nu(\theta) &= \sum_{\nu=2}^N \chi_\nu(\theta) + \chi\left(\frac{\theta}{2}\right) \\ &= \sum_{\nu=2}^N \chi_{\nu-1}\left(\frac{\theta}{2}\right) + \chi_0\left(\frac{\theta}{2}\right) \\ &= \sum_{\nu=0}^{N-1} \chi_\nu\left(\frac{\theta}{2}\right) \\ &= \dots = \chi\left(\frac{\theta}{2^N}\right) \end{aligned}$$

Donc  $\forall \theta \in \mathbb{R}^N$ , comme  $2^{-N} \xrightarrow{n \rightarrow \infty} 0$ , il existe un rang  $N_0 \in \mathbb{N}$  à partir duquel  $|2^{-N}\theta| \leq 1$  et donc  $\chi\left(\frac{\theta}{2^N}\right) = 1$  on a donc, pour tout  $N \geq N_0$  :

$$\sum_{\nu=0}^N \chi_\nu(\theta) = \chi\left(\frac{\theta}{2^N}\right) = 1 \xrightarrow{N \rightarrow \infty} 1$$

D'où  $\sum_{\nu} \chi_\nu \equiv 1$  sur  $\mathbb{R}^N$ .

Prenons maintenant  $a \in S_{0,1}^m$  et  $x \in K$  un compact de  $X$ . Nous allons établir des inégalités sur les différentielles. Par définition de  $S_{0,1}^m$  on a déjà :

$$|D_x^\beta D_\theta^\alpha a(x, \theta)| \leq M(1 + |\theta|)^{m-|\alpha|} \quad (5.9)$$

$\chi_\nu$  est  $C_0^\infty$  et est donc à décroissance rapide en l'infini, donc on peut trouver  $C_{\alpha,\nu}$  pour  $\alpha \in \mathbb{N}^N$  (a fortiori dépendant de  $\nu$ ) tel que :

$$|D^\alpha \chi_\nu(\theta)| \leq C_{\alpha,\nu}(1 + |\theta|)^{-|\alpha|} \quad (5.10)$$

Néanmoins, si on prend  $\mu, \nu > 0$  tel que  $\nu < \mu$  on a :

$$\begin{aligned} \chi_\mu(\theta) &= \chi_\nu(2^{\mu-\nu}\theta) \\ |D^\alpha \chi_\mu(\theta)| &= (2^{\mu-\nu})^{|\alpha|} |D^\alpha \chi_\nu(2^{\mu-\nu}\theta)| \\ &\leq C_{\alpha,\nu}(2^{\mu-\nu})^{|\alpha|} (1 + 2^{\mu-\nu}|\theta|)^{-|\alpha|} \\ &\leq C_{\alpha,\nu}(2^{\nu-\mu} + |\theta|)^{-|\alpha|} \\ &\leq C_{\alpha,\nu}(1 + |\theta|)^{-|\alpha|} \end{aligned}$$

Donc on peut prendre pour tout  $\nu > 0$ , la constante  $C_\alpha := \max(C_{\alpha,0}, C_{\alpha,1})$ . De (5.10) on déduit, pour  $\beta \in \mathbb{N}^n, \alpha \in \mathbb{N}^N$ , l'inégalité :

$$\begin{aligned} |D_x^\beta D_\theta^\alpha \chi_\nu(\theta) a(x, \theta)| &= |D_\theta^\alpha \chi_\nu(\theta) D_x^\beta a(x, \theta)| \\ &= \left| \sum_{\gamma \leq \alpha} D_\theta^\gamma \chi_\nu(\theta) D_\theta^{\alpha-\gamma} D_x^\beta a(x, \theta) \right| \\ &\leq \sum_{\gamma \leq \alpha} C_\gamma (1 + |\theta|)^{-|\gamma|} \times C_{\alpha,\gamma,\beta,K} (1 + |\theta|)^{m-|\alpha-\gamma|} \\ &= (1 + |\theta|)^{m-|\alpha|} \sum_{\gamma \leq \alpha} C_\gamma C_{\alpha,\gamma,\beta,K} \\ |D_x^\beta D_\theta^\alpha \chi_\nu(\theta) a(x, \theta)| &\leq \tilde{C}_{\alpha,\beta,K} (1 + |\theta|)^{m-|\alpha|} \end{aligned}$$

Ainsi,  $\chi_\nu a \in S_{0,1}^m$  pour tout  $\nu \geq 0$ , et comme  $\sum_{\nu} \chi_\nu \equiv 1$ ,  $\sum_{\nu} \chi_\nu a$  converge vers  $a$  dans  $S_{0,1}^m$ . Nous allons alors montrer que  $\sum_{\nu} I(\chi_\nu a u)$  converge vers  $I(au)$  ce qui permet de conclure la preuve du théorème.

Prenons donc  $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}^{n+N}$ ,  $2^{\nu-1} < 2^\nu |\theta| < 2^{\nu+1}$ , on a pour  $\nu > 0$  :

$$\begin{aligned} |D_{x,\theta}^\alpha a(x, 2^\nu \theta)| &\leq 2^{\nu|\alpha_2|} M(1 + 2^\nu \theta)^{m-|\alpha_2|} \\ &\leq M 2^{\nu|\alpha_2|} 2^{(\nu+2)(m-|\alpha_2|)} \\ &\leq M 2^{\nu m} 2^{-2|\alpha_2|} \leq CM 2^{\nu m} \end{aligned}$$

Et, pour  $\nu > 0$  on peut écrire, en faisant un changement de variables :

$$\begin{aligned} I_\phi(\chi_{\nu+1} a)(u) &= \int e^{i\phi(x,\theta)} \chi_{\nu+1}(\theta) a(x, \theta) u(x) dx d\theta \\ &= \int e^{i\phi(x,\theta)} \chi_1(2^{-\nu}\theta) a(x, \theta) dx d\theta \\ &= 2^{N\nu} \int e^{i2^\nu \phi(x,\theta)} \chi_1(\theta) a(x, 2^\nu \theta) u(x) dx d\theta \end{aligned}$$

Appliquons le lemme de phase non stationnaire 1 à la fonction  $C_0^N(K)$

$$(x, \theta) \in \mathbb{R}^n \times \mathbb{R}^N \mapsto \chi_1(\theta) a(x, 2^\nu \theta) u(x)$$

en notant  $K := \text{supp } u$  et  $\phi$  est une fonction de phase donc  $C^{N+1}(K)$  tel que  $d\phi \neq 0$  :

$$\begin{aligned} |I_\phi(\chi_{\nu+1} a)| &= 2^{N\nu} \left| \int e^{i2^\nu \phi(x, \theta)} \chi_1(\theta) a(x, 2^\nu \theta) u(x) dx d\theta \right| \\ &\leq 2^{N\nu-k\nu} C \sum_{|\alpha| \leq k} \sup |D^\alpha(\chi_1(\theta) a(x, 2^\nu \theta) u(x))| |\phi'|^{|\alpha|-2k} \end{aligned}$$

et :

$$\begin{aligned} |D^\alpha(\chi_1(\theta) a(x, 2^\nu \theta) u(x))| &= \left| \sum_{\beta \leq \alpha} \sum_{\gamma \leq \beta} D^\beta u(x) D^{\alpha-\beta} a(x, 2^\nu \theta) D^{\alpha-\beta-\gamma} \chi_1(\theta) \right| \\ &\leq \sum_{\beta \leq \alpha} |D^\beta u(x)| |D^{\alpha-\beta} a(x, 2^\nu \theta)| |D^{\alpha-\beta-\gamma} \chi_1(\theta)| \\ &\leq C \sum_{\beta \leq \alpha} |D^\beta u(x)| \times CM 2^{\nu m} \end{aligned} \quad (5.11)$$

Donc :

$$\begin{aligned} |I_\phi(\chi_{\nu+1} a)| &\leq 2^{N\nu-k\nu+\nu m} \tilde{C} M \sum_{|\alpha| \leq k} \sup |D^\alpha u| \\ |I_\phi(\chi_{\nu+1} a)| &\leq C_\nu \|u\|_{C^k} \end{aligned} \quad (5.12)$$

où la constante  $C_\nu$  dépend de  $\nu, m, k, K$  et non de  $u$ , tel que  $\sum_\nu C_\nu$  est sommable si  $k > N + m$ .

Ainsi, par comparaison avec le terme général d'une série convergente, dès lors que  $k > N + m$ , on a que la série  $\sum_\nu |I_\phi(\chi_\nu a u)|$  converge et on peut alors intervertir la somme et l'intégrale pour obtenir :

$$I_\phi(a u) = \sum_\nu I_\phi(\chi_\nu a u)$$

et on retrouve par comparaison :

$$|I_\phi(a u)| \leq \sum_\nu C_\nu \|u\|_{C^k}$$

Donc  $I_\phi(a)$  définit pour tout  $a \in S_{0,1}^m$  une distribution d'ordre au plus  $k$  si  $k > N + m$ , ce qui conclue la preuve du théorème.  $\square$

**Exemple 5.3.1.** Un exemple important d'intégrale oscillante est :

$$\int e^{ix \cdot \theta} d\theta, \quad \theta \in \mathbb{R}^n, x \in \mathbb{R}^n$$

Prenons  $u \in C_0^\infty(\mathbb{R}^n)$ , et considérons  $\rho \in C_0^\infty(X)$  un noyau de convolution tel que l'on ait  $\text{supp } \rho \subset [-1; 1]$  et  $\int \rho = 1$ , pour tout  $t > 0$  :

$$\begin{aligned} \int e^{ix \cdot \theta} \rho\left(\frac{\theta}{t}\right) u(x) dx d\theta &= t^n \int \hat{\rho}(-tx) u(x) dx \\ &= \int \hat{\chi}(-x) u\left(\frac{x}{t}\right) dx \\ &\xrightarrow{t \rightarrow 0} u(0) \int \hat{\rho}(-x) dx \\ &= u(0) \times \int e^{ix \cdot 0} \hat{\rho}(-x) dx \\ &= u(0) \times (2\pi)^n \rho(0) \\ &= u(0) (2\pi)^n \end{aligned}$$

Donc, la distribution définie par l'intégrale oscillante est la limite de la double intégrale que l'on vient de calculer et on a donc la formule :

$$\int e^{ix\theta} d\theta = \delta_0(2\pi)^n$$

### 5.3.3. Front d'onde d'une intégrale oscillante

Nous allons maintenant pouvoir discuter du front d'onde des intégrales oscillantes. Nous allons nous appuyer sur les résultats donnés sur le front d'onde dans la partie 5.2.3, et notamment la remarque 5.2.3.

#### Théorème 5.3.2 -

Soit  $A := I_\phi(a) = \int e^{i\phi(\cdot, \theta)} a(\cdot, \theta) d\theta$  la distribution définie sous les hypothèses du théorème 1. On a alors :

$$WF(I_\phi(a)) \subset \{(x, \nabla_x \phi(x, \theta)) : (x, \theta) \in F \text{ et } \nabla_\theta \phi(x, \theta) = 0\}$$

*Démonstration.* Soit  $\psi \in C_0^\infty(X)$ ,  $\varphi \in \mathcal{D}(X)$ , on remarque que :

$$\begin{aligned} \langle \widehat{\psi A}(\xi), \varphi(\xi) \rangle &= \langle \psi A(x), \widehat{\varphi}(x) \rangle \\ &= \langle A, \psi \widehat{\varphi} \rangle \\ &= \int e^{i\phi(x, \theta)} a(x, \theta) \psi(x) \int e^{-ix \cdot \xi} \varphi(\xi) d\xi dx d\theta \\ &= \int \left( \int e^{i\phi(x, \theta) - ix \cdot \xi} a(x, \theta) \psi(x) dx d\theta \right) \varphi(\xi) d\xi \end{aligned}$$

Donc :

$$\widehat{\psi A} \stackrel{\mathcal{D}'(X)}{=} \int e^{i\phi(x, \theta) - ix \cdot \xi} a(x, \theta) \psi(x) dx d\theta$$

Soit  $V \subset \mathbb{R}^n$  un cône fermé tel que pour  $(x, \theta) \in F$  avec  $x \in \text{supp } \psi$ , et  $\nabla_\theta \phi(x, \theta) = 0$  alors  $\nabla_x \phi(x, \theta) \notin V$ . On considère  $(x, \theta) \in F$ , avec  $x \in \text{supp } \psi$  et  $\xi \in V$ , on cherche à montrer qu'il existe  $c > 0$  tel que :

$$|\xi - \nabla_x \phi(x, \theta)| + |\theta| |\nabla_\theta \phi(x, \theta)| \geq c(|\xi| + |\theta|) \quad (5.13)$$

- Si  $\theta = 0$  : alors pour tout  $t > 0$ , on a  $\phi(x, t\theta) = t\phi(x, \theta) = \phi(x, 0)$  donc par continuité, on déduit quand  $t \rightarrow 0$  que  $\phi(x, 0) = 0$ , donc la fonction  $x \mapsto \phi(x, 0)$  est constante, d'où  $\nabla_x \phi(x, 0) = 0$  et  $|\xi - \nabla_x \phi(x, 0)| = |\xi|$ . On a donc finalement, dans ce cas :

$$|\xi - \nabla_x \phi(x, \theta)| + |\theta| |\nabla_\theta \phi(x, \theta)| \geq |\xi| = (|\xi| + |\theta|)$$

- Si  $\theta \neq 0$  : on va dans un premier temps supposer que  $|\xi| + |\theta| = 1$ . Alors, on peut remarquer que  $\xi \in V \cap \overline{B}_n(0, 1)$  et  $(x, \theta) \in F \cap (\text{supp } \psi \times \overline{B}_N(0, 1))$ , i.e. la fonction  $(x, \theta, \xi) \mapsto |\xi - \nabla_x \phi(x, \theta)| + |\theta| |\nabla_\theta \phi(x, \theta)|$  est continue sur un compact. Donc pour montrer l'inégalité (5.13), il suffit de vérifier que cette dernière ne s'annule jamais.

- Si  $\nabla_\theta \phi(x, \theta) \neq 0$  alors  $|\theta| |\nabla_\theta \phi(x, \theta)| > 0$  et donc

$$|\xi - \nabla_x \phi(x, \theta)| + |\theta| |\nabla_\theta \phi(x, \theta)| > 0$$

- Si  $\nabla_\theta \phi(x, \theta) = 0$ , alors comme  $\xi \in V$  on a forcément  $\xi \notin \{\nabla_x \phi(x, \theta) : (x, \theta) \in F, x \in \text{supp } \psi, \nabla_\theta \phi(x, \theta) = 0\}$  et donc  $\xi \neq \nabla_x \phi(x, \theta)$ , d'où  $|\xi - \nabla_x \phi(x, \theta)| > 0$  et :

$$|\xi - \nabla_x \phi(x, \theta)| + |\theta| |\nabla_\theta \phi(x, \theta)| > 0$$

Dans tout les cas, on a bien :

$$|\xi - \nabla_x \phi(x, \theta)| + |\theta| |\nabla_\theta \phi(x, \theta)| > 0$$

et donc par continuité sur un compact, il existe une constante  $c > 0$  tel que :

$$|\xi - \nabla_x \phi(x, \theta)| + |\theta| |\nabla_\theta \phi(x, \theta)| \geq c$$

Si on considère maintenant  $|\xi| + |\theta|$  quelconque -non nul, sans quoi le résultat est trivial- et on pose  $\xi_1 = \frac{\xi}{|\xi| + |\theta|}$  et  $\theta_1 = \frac{\theta}{|\xi| + |\theta|}$  de sorte que :

$$|\xi_1| + |\theta_1| = \frac{|\xi| + |\theta|}{|\xi| + |\theta|} = 1$$

La fonction  $\phi$  est positivement homogène en  $\theta$  donc  $\nabla_x \phi$  est positivement homogène en  $\theta$  tandis que  $\nabla_\theta \phi$  est homogène en  $\theta$  de degré 0 ; on a alors :

$$\begin{aligned} |\xi_1 - \nabla_x \phi(x, \theta_1)| + |\theta_1| |\nabla_\theta \phi(x, \theta_1)| &\geq c \\ \left| \frac{\xi}{|\xi| + |\theta|} - \frac{1}{|\xi| + |\theta|} \nabla_x \phi(x, \theta) \right| + \frac{|\theta|}{|\xi| + |\theta|} |\nabla_\theta \phi(x, \theta)| &\geq c \\ |\xi - \nabla_x \phi(x, \theta)| + |\theta| |\nabla_\theta \phi(x, \theta)| &\geq c(|\xi| + |\theta|) \end{aligned}$$

On a donc montré l'équation (5.13).

Si on reprend la démonstration du théorème 1, on peut écrire, pour la même partition  $(\chi_\nu)$ , que :

$$\widehat{\psi A}(\xi) = \sum_\nu \int e^{i(\phi(x, \theta) - x \cdot \xi)} \psi(x) a(x, \theta) \chi_\nu(\theta) dx d\theta$$

Et en effectuant le même changement de variable on obtient également :

$$\begin{aligned} \int e^{i(\phi(x, \theta) - x \cdot \xi)} \chi_\nu(\theta) a(x, \theta) \psi(x) dx d\theta \\ = 2^{N(\nu-1)} \int e^{i(2^{\nu-1} \phi(x, \theta) - x \cdot \xi)} \chi_1(\theta) \psi(x) a(x, 2^{\nu-1} \theta) dx d\theta \end{aligned}$$

On pose  $\Phi : (x, \theta) \mapsto \frac{R\phi(x, \theta) - x \cdot \xi}{R + |\xi|}$  où  $R = 2^{\nu-1}$  et  $\xi \in V$ , on a alors :

$$\begin{aligned} \nabla_x \Phi(x, \theta) &= \frac{R \nabla_x \phi(x, \theta) - \xi}{R + |\xi|} \\ \nabla_\theta \Phi(x, \theta) &= \frac{R \nabla_\theta \phi(x, \theta)}{R + |\xi|} \end{aligned}$$

et donc d'après (5.13) :

$$\begin{aligned} |\nabla_x \Phi(x, \theta)| + |\nabla_\theta \Phi(x, \theta)| &= \frac{1}{R + |\xi|} (|\xi - R \nabla_x \phi(x, \theta)| + R |\nabla_\theta \phi(x, \theta)|), \\ &\geq c \frac{R|\theta| + \xi}{R + |\xi|} \geq \tilde{c} \end{aligned}$$

et, en réutilisant (5.11) :

$$|D_\theta^\alpha D_x^\beta (\psi(x) \chi_1(\theta) a(x, R\theta))| \leq C M R^m \times \|\psi\|_{C^{|\beta|}} \leq \tilde{C}_{\alpha, \beta} M R^m$$

Ainsi, si on applique le lemme 1, avec  $\Phi$  ci-dessus et  $|\nabla\Phi| = |\nabla_x\Phi| + |\nabla_\theta\Phi| \neq 0$  on retrouve pour  $k \in \mathbb{N}$  :

$$\begin{aligned} \frac{R^N}{(R + |\xi|)^k} \times (R + |\xi|)^k \left| \int e^{i(R+|\xi|)\Phi(x,\theta)} \chi_1(\theta) \psi(x) a(x, R\theta) dx d\theta \right| \\ \leq R^N (R + |\xi|)^{-k} C \sum_{|\alpha| \leq k} \sup |D^\alpha(\psi(x) \chi_1(\theta) a(x, R\theta))| |\nabla\Phi|^{|\alpha|-2k} \\ \leq CR^N (R + |\xi|)^{-k} \times \tilde{C} R^m \times \tilde{c}' \\ \leq CR^{m+N} (R + |\xi|)^{-k} \end{aligned}$$

Si on prend  $k$  suffisamment grand et notamment plus grand que  $N + m + 1$  on a :

$$\begin{aligned} \left( \frac{R}{|\xi|} \right)^{N+m+1} &\leq \left( \frac{R}{|\xi|} + 1 \right)^{N+m+1} \\ &\leq \left( \frac{R}{|\xi|} + 1 \right)^k \\ \Rightarrow \frac{R^{m+N}}{(R + |\xi|)^k} &\leq \frac{|\xi|^{m+N+1-k}}{R} \end{aligned}$$

Finalement :

$$\int e^{i(\phi(x,\theta)-x \cdot \xi)} \chi_\nu(\theta) a(x, \theta) \psi(x) dx d\theta \leq CR^{-1} |\xi|^{m+N+1-k}$$

où la constante  $C$  ne dépend pas de  $\nu$ , de sorte que  $\sum_\nu R^{-1} = 2$  et donc par comparaison avec le terme général d'une série convergente, la série :

$$\sum_\nu \left| \int e^{i(\phi(x,\theta)-x \cdot \xi)} \chi_\nu(\theta) a(x, \theta) \psi(x) dx d\theta \right|$$

converge et en intervertissant ensuite la somme et l'intégrale on retrouve que :

$$|\widehat{\psi A}(\xi)| \leq C |\xi|^{m+N+1-k}$$

pour tout  $k$  assez grand et tout  $\xi \in V$ . Donc la fonction  $\widehat{\psi A}$  décroît rapidement sur tout cône n'intersectant pas  $\{\nabla_x \phi(x, \theta) : x \in \text{supp } \psi, (x, \theta) \in F, \nabla_\theta \phi(x, \theta) = 0\}$ .

Donc  $A$  n'est pas microlocalement  $C^\infty$  en tout point  $(x, \xi)$  tel que  $\xi = \nabla_x \phi(x, \theta)$  où  $(x, \theta) \in F$  tel que  $\nabla_\theta \phi(x, \theta) = 0$ , i.e. finalement :

$$WF(I_\phi(a)) \subset \{(x, \nabla_x \phi(x, \theta)) : (x, \theta) \in F \text{ et } \nabla_\theta \phi(x, \theta) = 0\}$$

□

---

## 5.4. Du produit de distributions

---

Le but de cette partie va être d'étudier le produit de distributions en considérant le point de vue microlocal, et de chercher à savoir quand il peut être défini.

### 5.4.1. Généralités

Dans toute la suite, on considérera un ouvert  $\Omega$  de  $\mathbb{R}^n$ .

#### Rappels

Pour entamer l'étude du produit de distributions, il est naturel de commencer par rappeler le cas très simple du produit par une fonction  $C^\infty$  vu précédemment.

#### Définition 5.4.1 -

Soit  $u \in \mathcal{D}'(\Omega)$  et  $\phi \in C^\infty(\Omega)$ . La distribution  $\phi u$  est bien définie par :

$$\forall \varphi \in C_0^\infty(\Omega), \langle \phi u, \varphi \rangle = \langle u, \phi \varphi \rangle$$

Il n'est cependant pas possible de définir le produit de distributions plus générales, même si elles sont élémentaires, si l'on veut que ce produit conserve quelque propriété judicieuse. Prenons l'exemple du produit de  $\delta_0$  par la fonction  $H$  de Heaviside dans  $\mathcal{D}'(\mathbb{R})$ . Il ne pourra alors pas y avoir continuité du produit. En effet, si l'on considère la suite des fonctions  $C^\infty (f_n)_{n \geq 1}$  telles que  $f_n$  soit nulle sur  $]-\infty, -1/n[$  et vaille 1 sur  $]1/n, +\infty[$ , on obtient que cette suite converge vers  $H$  au sens des distributions. Il serait alors raisonnable de définir le produit  $\delta H$  comme la limite de  $\delta f_n$ . Or  $\delta f_n = \delta f_n(0)$  et il est tout à fait possible de choisir les  $f_n$  de sorte que  $f_n(0)$  n'ai pas de limite. De plus, même en renonçant à la propriété de continuité, il n'est pas possible de définir un tel produit en conservant au moins des propriétés algébriques, comme la formule de Leibniz. On a en effet que pour tout  $n$ ,  $H^n = H$  et donc on devrait obtenir :

$$\delta = H' = (H^n)' = nH^{n-1}H' = n(H\delta)$$

D'où

$$H\delta = \frac{1}{n}\delta, \forall n \in \mathbb{N}^*$$

Ce produit aurait alors une infinité de valeurs différentes, ce qui n'est pas cohérent.

Ces quelques exemples indiquent qu'il n'y a pas de manière naturelle est immédiate pour définir un produit de distributions dans le cas général. On ne sait pas raisonnablement multiplier deux distributions quelconques entre elles.

#### Quelques cas particuliers

Il existe néanmoins des cas où la singularité d'un des deux termes peut être compensée par la régularité de l'autre, et le produit défini. On peut par exemple définir  $uv$  dans les cas suivants :

- si  $u \in L_{loc}^1$  et  $v \in L_{loc}^\infty$
- si  $u \in L_{loc}^2$  et  $v \in L_{loc}^2$
- plus généralement si  $u \in L_{loc}^p$  et  $v \in L_{loc}^q$ , avec  $\frac{1}{p} + \frac{1}{q} = 1$

car alors  $uv$  sera  $L_{loc}^1$  par Hölder.

On peut également définir le produit  $uv$  lorsque  $u$  est une distribution d'ordre  $\leq p$  et que  $v \in C^p(\Omega)$ . Pour définir un tel produit, on commence par prolonger  $u$ , qui est une application linéaire sur  $C_0^\infty(\Omega)$ , à  $C_0^p(\Omega)$ . Soit  $\varphi \in C_0^p(\Omega)$  de support  $K$ . Soit  $K' \subset \Omega$  un voisinage compact de  $K$ . Par densité de  $C_0^\infty(\Omega)$ ,



on peut trouver  $(\varphi_j)_j \in (C_0^\infty(\Omega))^\mathbb{N}$  telle que  $\forall j, \text{supp}(\varphi_j) \subset K'$  et  $\varphi_j \rightarrow \varphi$  ainsi que ses dérivées jusqu'à l'ordre  $p$ . Or,  $u$  étant d'ordre  $p$ , on a :

$$|\langle u, \varphi_j \rangle| \leq C. \sup_{x \in K', |\alpha| \leq p} |\partial^\alpha \varphi_j(x)|$$

On peut ainsi définir  $\langle u, \varphi \rangle$  comme  $\lim_{j \rightarrow \infty} \langle u, \varphi_j \rangle$  par le théorème de prolongement de la convergence. Comme  $\forall \varphi \in C_0^\infty(\Omega), v\varphi \in C_0^p(\Omega)$ , nous pouvons poser :

$$\langle uv, \varphi \rangle = \langle u, v\varphi \rangle$$

et on vérifie à l'aide la formule de Leibniz que la forme linéaire ainsi obtenue est bien une distribution.

Le produit de deux distributions peut aussi être défini grâce au support singulier précédemment défini. Avec cette notion, on peut localiser les singularités, ce qu'on ne faisait pas avec les exemples précédents.

Soit  $u, v \in \mathcal{D}'(\Omega)$  vérifiant

$$\text{supp sing}(u) \cap \text{supp sing}(v) = \emptyset.$$

On considère le recouvrement de  $\Omega$  par les ouverts :

$$U_1 = \Omega \setminus \text{supp sing}(u) \text{ et } U_2 = \Omega \setminus \text{supp sing}(v).$$

Avec une partition de l'unité, toute fonction  $\varphi \in C_0^\infty(\Omega)$  peut s'écrire

$$\varphi = \varphi_1 + \varphi_2$$

avec  $\varphi_1 \in C_0^\infty(U_1)$  et  $\varphi_2 \in C_0^\infty(U_2)$ . On pose alors :

$$\langle uv, \varphi \rangle = \langle u, \varphi_2 v \rangle + \langle v, \varphi_1 u \rangle$$

ce qui est bien défini, car  $\varphi_1 u$  et  $\varphi_2 v$  sont  $C_0^\infty(\Omega)$  et le résultat ne dépend pas de la partition de l'unité choisie.

### Produit tensoriel

Il est cependant toujours possible d'avoir une sorte de "produit" pour deux distributions grâce au produit tensoriel. Commençons par définir le produit tensoriel de deux fonctions.

Définition 5.4.2 -

Soit  $f$  et  $g$  deux fonctions définies sur des ouverts de  $\mathbb{R}^p$  et  $\mathbb{R}^q$  respectivement.  
Le *produit tensoriel*  $f \otimes g$  est la fonction définie dans l'ouvert produit de  $\mathbb{R}^{p+q}$   
par

$$f \otimes g(x, y) = f(x)g(y)$$

Ce produit tensoriel s'étend aux distributions sans aucune condition.

Définition 5.4.3 -

Soient  $\Omega_1 \subset \mathbb{R}^p$  et  $\Omega_2 \subset \mathbb{R}^q$  deux ouverts. Soit  $u \in \mathcal{D}'(\Omega_1)$  et  $v \in \mathcal{D}'(\Omega_2)$ . Le *produit tensoriel*  $u \otimes v \in \mathcal{D}'(\Omega_1 \times \Omega_2)$  est défini par :

$$\forall \varphi \in C_0^\infty(\Omega_1 \times \Omega_2), \langle u \otimes v, \varphi \rangle = \langle u_x, \langle v_y, \varphi_{x,y} \rangle \rangle = \langle v_y, \langle u_x, \varphi_{x,y} \rangle \rangle$$

*Remarque.* Les lettres muettes  $x$  et  $y$  en indice servent à indiquer contre quelle variable tape chacune des distributions.

*Remarque.* Le produit tensoriel est ainsi bien défini grâce au théorème de Fubini et au théorème de dérivation sous le crochet. Ce dernier nous permet en effet d'affirmer que les applications  $x \mapsto \langle v_y, \varphi(x, y) \rangle$  et  $y \mapsto \langle u_x, \varphi(x, y) \rangle$  sont bien  $C_0^\infty(\Omega_1)$  et  $C_0^\infty(\Omega_2)$  respectivement.

**Exemples 5.4.1.** Pour  $\Omega_1 = \Omega_2 = \mathbb{R}$  :

- la distribution  $\delta_0 \otimes \delta_0$  n'est autre que le Dirac  $\delta_{(0,0)}$  de  $\mathbb{R}^2$ .
- la distribution  $\mathbf{1}_{\mathbb{R}} \otimes \delta_0$  est définie par :

$$\forall \varphi \in C_0^\infty(\mathbb{R}^2), \langle \mathbf{1}_{\mathbb{R}} \otimes \delta_0, \varphi \rangle = \int_{\mathbb{R}} \varphi(x, 0) dx.$$

Il s'agit de la mesure de densité linéique 1 portée par l'axe des  $x$ .

*Remarque.* Pour des fonctions  $g$  et  $f$  définies sur un ensemble quelconque  $X$ , on peut définir le produit  $fg$  de manière détournée en utilisant le produit tensoriel. En effet,  $f \otimes g$  est la fonction définie sur  $X \times X$  par  $(x, y) \mapsto f(x)g(y)$ . Si l'on considère sa restriction à la diagonale, on obtient une fonction définie sur  $X$  par  $x \mapsto f(x)g(x)$ , et on retrouve le produit  $fg$ . C'est de cette idée là que l'on va s'inspirer pour définir le produit de deux distributions dans un cas plus général. Viens alors la nécessité de prendre en compte une description plus raffinée des singularités.

### 5.4.2. Front d'onde et produit

Grâce au front d'onde, nous allons maintenant décrire les singularités des distributions afin de définir leur produit. Pour ce faire, nous allons étendre continûment le cas  $C^\infty$ , c'est pourquoi il est nécessaire de définir une topologie sur les distributions munies de leur front d'onde.

#### Notions de convergence

Dans toute la suite, on notera  $X$  un ouvert de  $\mathbb{R}^n$  et  $\Gamma$  un cône fermé de  $X \times (\mathbb{R}^n \setminus 0)$ .

*Définition 5.4.4 -*

On pose :

$$\mathcal{D}'_\Gamma(X) = \{u \in \mathcal{D}'(X), WF(u) \subset \Gamma\}.$$

*Proposition 5.4.0 -*

Soit  $u \in \mathcal{D}'(X)$ . Alors  $u \in \mathcal{D}'_\Gamma(X)$  si, et seulement si pour tout  $\phi \in C_0^\infty(X)$  et pour tout cône fermé  $V \subset \mathbb{R}^n$  tels que :

$$\Gamma \cap (supp \phi \times V) = \emptyset, \quad (5.14)$$

on a :

$$\forall N \in \mathbb{N}^*, \sup_V |\xi|^N |\widehat{\phi u}(\xi)| < \infty. \quad (5.15)$$

*Démonstration.* Si  $u \in \mathcal{D}'_\Gamma(X)$ , soit  $(x, \xi) \notin \Gamma$ , on prend  $\phi \in C_0^\infty(X)$  tel que  $\phi(x) \neq 0$  et  $V$  un cône fermé tel que  $\xi \in \mathring{V}$ . Quitte à rétrécir  $supp \phi$  et  $V$ , on peut supposer  $\Gamma \cap (supp \phi \times V) = \emptyset$ . Alors

$$\forall N \in \mathbb{N}^*, \sup_V |\xi|^N |\widehat{\phi u}(\xi)| < \infty$$

et donc

$$(x, \xi) \notin WF(u)$$

par la remarque (5.2.3). Ceci montre que  $u \in \mathcal{D}'_\Gamma(X)$ .

Si  $u \in \mathcal{D}'_\Gamma(X)$ , soit  $\phi \in C_0^\infty(X)$  et un cône fermé  $V \subset \mathbb{R}^n$  tels que :

$$\Gamma \cap (\text{supp } \phi \times V) = \emptyset.$$

Soit  $\xi \in \Sigma(\phi u)$ . Alors, par (2), il existe  $x \in \text{supp } \phi$  tel que  $(x, \xi) \in WF(\phi u) \stackrel{(3)}{\subset} WF(u) \subset \Gamma$ , donc  $\xi \notin V$  par (5.14). D'où (5.15).  $\square$

#### Définition 5.4.5 -

Pour toute suite  $(u_j)$  de  $\mathcal{D}'_\Gamma(X)$  et  $u \in \mathcal{D}'_\Gamma(X)$ , on dira que  $u_j$  converge vers  $u$  dans  $\mathcal{D}'_\Gamma(X)$  si :

(i)  $u_j \rightarrow u$ , dans  $\mathcal{D}'(X)$

(ii)  $\forall N \in \mathbb{N}^*$ ,  $\phi \in C_0^\infty(X)$  et  $V \subset \mathbb{R}^n$  un cône fermé tels que  $\Gamma \cap (\text{supp } \phi \times V) = \emptyset$ ,

$$\sup_V |\xi|^N \left| \widehat{\phi u}(\xi) - \widehat{\phi u_j}(\xi) \right| \xrightarrow{j \rightarrow \infty} 0$$

*Remarque.* Puisque (i) implique que  $\widehat{\phi u_j} \rightarrow \widehat{\phi u}$  uniformément sur tout compact et que  $N$  est quelconque dans (ii), on peut remplacer (ii) par :

(ii)'  $\forall N \in \mathbb{N}^*$ ,  $\phi \in C_0^\infty(X)$  et  $V \subset \mathbb{R}^n$  un cône fermé tels que  $\Gamma \cap (\text{supp } \phi \times V) = \emptyset$ ,

$$\sup_V |\xi|^N |\widehat{\phi u_j}(\xi)| < \infty.$$

#### Théorème 5.4.1 -

Pour tout  $u \in \mathcal{D}'_\Gamma(X)$ , il existe une suite  $(u_j)_{j \in \mathbb{N}} \in (C_0^\infty(X))^{\mathbb{N}}$  telle que

$$u_j \rightarrow u \text{ dans } \mathcal{D}'_\Gamma(X).$$

*Démonstration.* Comme dans le cas des distributions, cette propriété s'obtient en construisant la suite  $u_j$  par convolution à partir de fonctions plateaux. On en trouvera une preuve explicite dans les travaux d'Hörmander [1].  $\square$

#### Image réciproque

Nous allons maintenant nous intéresser au théorème principal qui nous aidera à définir le produit de distribution, puisqu'il permettra de généraliser la notion de restriction.

#### Définition 5.4.6 -

Soient  $X$  un ouvert de  $\mathbb{R}^m$  et  $Y$  un ouvert de  $\mathbb{R}^n$ . Soit  $f \in C^\infty(X, Y)$ . L'ensemble des normales à  $f$  est

$$N_f \stackrel{\text{def}}{=} \{(f(x), \eta) \in Y \times \mathbb{R}^n; {}^t f'(x)\eta = 0\},$$

où  $f'(x)$  désigne la différentielle de  $f$  en  $x$ .

#### Théorème 5.4.2 -

Soient  $X$  un ouvert de  $\mathbb{R}^m$  et  $Y$  un ouvert de  $\mathbb{R}^n$ . Soit  $f \in C^\infty(X, Y)$  et  $u \in \mathcal{D}'(Y)$  telle que

$$N_f \cap WF(u) = \emptyset.$$

Alors l'image réciproque  $f^*u$  peut être définie d'une et une seule manière pour que  $f^*u = u \circ f$  lorsque  $u \in C^\infty(Y)$  et que pour tout  $\Gamma$  cône fermé de  $Y \times (\mathbb{R}^n \setminus 0)$  vérifiant  $N_f \cap \Gamma = \emptyset$ , l'application  $f^* : \mathcal{D}'_\Gamma(Y) \rightarrow \mathcal{D}'_{f^*\Gamma}(X)$  est continue, où l'on note :

$$f^*\Gamma = \{(x, {}^t f'(x)\eta) ; (f(x), \eta) \in \Gamma\}.$$

Corollaire 5.4.2 -

On a en particulier que pour tout  $u \in \mathcal{D}'(Y)$  vérifiant  $N_f \cap WF(u) = \emptyset$ ,

$$WF(f^*u) \subset f^*WF(u) \quad (5.16)$$

*Démonstration.* En effet, pour tout cône fermé  $\Gamma \in Y \times (\mathbb{R}^n \setminus 0)$  tel que  $WF(u) \subset \Gamma$ , on a d'après le théorème 2 que  $f^*(\mathcal{D}'_\Gamma(Y)) \subset \mathcal{D}'_{f^*\Gamma}(X)$ , d'où  $f^*(u) \in \mathcal{D}'_{f^*\Gamma}(X)$ . Donc, par définition,  $WF(f^*u) \subset f^*\Gamma$ . Si on prend  $\Gamma = WF(u)$ , on a en particulier que  $WF(f^*u) \subset f^*WF(u)$ .  $\square$

Pour démontrer le théorème 2, nous allons nous servir du lemme suivant, qui a été démontré dans la partie 5.3.1.

Lemme 5.4.2 -

Soit  $K \subset \mathbb{R}^n$  un compact, et  $X$  un voisinage ouvert de  $K$ . Soit  $k \geq 0$ .

Si  $u \in C_0^k(K)$ ,  $f \in C^{k+1}(K)$  à valeurs réelles sur  $X$  tel que  $|f'| \neq 0$  sur  $K$ . Alors, il existe une constante  $C$ , bornée lorsque  $f$  reste bornée dans  $C^{k+1}(X)$  tel que :

$$\forall \omega > 0, \quad \omega^k \left| \int u(x) e^{i\omega f(x)} dx \right| \leq C \sum_{|\alpha| \leq k} \sup |D^\alpha u| |f'|^{|\alpha| - 2k}$$

*Démonstration.* Commençons d'abord par définir l'image réciproque comme une limite dans  $\mathcal{D}'(X)$ . Soit  $u \in \mathcal{D}'(Y)$  vérifiant  $N_f \cap WF(u) = \emptyset$ . Soit  $\chi \in C_0^\infty(X)$ . Fixons un fermé conique  $\Gamma \subset Y \times (\mathbb{R}^n \setminus 0)$  tel que  $WF(u) \subset \Gamma$  et  $\Gamma \cap N_f = \emptyset$ .

Soit  $v \in C_0^\infty(Y)$ . On peut déjà remarquer que

$$\langle f^*v, \chi \rangle = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{v}(\eta) I_\chi(\eta) d\eta, \quad (5.17)$$

$$I_\chi(\eta) = \int_{\mathbb{R}^m} \chi(x) e^{i\langle f(x), \eta \rangle} dx. \quad (5.18)$$

En effet, on a

$$\begin{aligned} \langle f^*v, \chi \rangle &= \langle v \circ f, \chi \rangle \\ &= \langle (\mathcal{F}^{-1} \circ \mathcal{F}(v)) \circ f, \chi \rangle \\ &= \left\langle \left[ \frac{1}{(2\pi)^n} \sigma \circ \mathcal{F}(\widehat{v}) \right] \circ f, \chi \right\rangle, \end{aligned}$$

Il suit

$$\begin{aligned}\langle f^*v, \chi \rangle &= \left\langle \frac{1}{(2\pi)^n} \left[ \sigma \circ \left( \int_{\mathbb{R}^n} \widehat{v}(\eta) e^{-i\langle \cdot, \eta \rangle} d\eta \right) \right] \circ f, \chi \right\rangle \\ &= \frac{1}{(2\pi)^n} \left\langle \left[ \int_{\mathbb{R}^n} \widehat{v}(\eta) e^{i\langle \cdot, \eta \rangle} d\eta \right] \circ f, \chi \right\rangle \\ &= \frac{1}{(2\pi)^n} \left\langle \int_{\mathbb{R}^n} \widehat{v}(\eta) e^{i\langle f(\cdot), \eta \rangle} d\eta, \chi \right\rangle\end{aligned}$$

Or, pour tout  $K$  compact de  $\mathbb{R}^m$ ,

$$\int_K \left| \int_{\mathbb{R}^n} \widehat{v}(\eta) e^{i\langle f(x), \eta \rangle} d\eta \right| dx \leq \int_K \int_{\mathbb{R}^n} |\widehat{v}(\eta)| d\eta dx \leq |K| \cdot \int_{\mathbb{R}^n} |\widehat{v}(\eta)| d\eta < \infty$$

car  $v \in C_0^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ , donc  $\widehat{v} \in \mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$ .

On a donc

$$x \mapsto \int_{\mathbb{R}^n} \widehat{v}(\eta) e^{i\langle f(x), \eta \rangle} d\eta \in L_{loc}^1(\mathbb{R}^m) \text{ et } \langle f^*v, \chi \rangle = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^m} \left( \int_{\mathbb{R}^n} \widehat{v}(\eta) e^{i\langle f(x), \eta \rangle} d\eta \right) \chi(x) dx.$$

De même,

$$\int_{\mathbb{R}^m} \int_{\mathbb{R}^n} |\widehat{v}(\eta) e^{i\langle f(x), \eta \rangle} \chi(x)| d\eta dx \leq |\text{supp } \chi| \|\chi\|_\infty \int_{\mathbb{R}^n} |\widehat{v}(\eta)| d\eta < \infty$$

Par Fubini, on obtient :

$$\begin{aligned}\langle f^*v, \chi \rangle &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{v}(\eta) \left( \int_{\mathbb{R}^m} \chi(x) e^{i\langle f(x), \eta \rangle} dx \right) d\eta \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{v}(\eta) I_\chi(\eta) d\eta.\end{aligned}$$

Soit maintenant  $x_0 \in X$ , et posons  $y_0 = f(x_0)$ ,  $\Gamma_{y_0} = \{\eta ; (y_0, \eta) \in \Gamma\}$ . On pose :

- $V \subset \mathbb{R}^n \setminus \{0\}$  un voisinage conique fermé de  $\Gamma_{y_0}$  tel que  ${}^t f'(x_0)\eta \neq 0$ ,  $\forall \eta \in V$ ,
- $Y_0 \subset Y$  un voisinage compact de  $y_0$  tel que  $V$  est un voisinage de  $\Gamma_y$ ,  $\forall y \in Y_0$  (un tel  $Y_0$  existe car  $\Gamma$  est fermé),
- $X_0 \in X$  un voisinage compact de  $x_0$  tel que  $f(X_0) \subset Y_0^\circ$  et  ${}^t f'(x)\eta \neq 0$ ,  $\forall x \in X_0$ ,  $\eta \in V$ .

Soit  $\phi \in C_0^\infty(Y_0)$  tel que  $\phi|_{f(X_0)} \equiv 1$ . Alors (5.17) est toujours valable pour  $\chi \in C_0^\infty(X_0)$  et  $v \in C^\infty(Y)$  si  $v$  est remplacé par  $\phi v$ . En effet,  $\phi v \in C_0^\infty(Y)$  et ainsi

$$\langle f^*(\phi v), \chi \rangle = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{\phi v}(\eta) I_\chi(\eta) d\eta.$$

De plus,

$$\begin{aligned}\langle f^*(\phi v), \chi \rangle &= \int_{\mathbb{R}^m} (\phi v) \circ f(x) \chi(x) dx \\ &= \int_{X_0} (\phi \circ f(x)) (v \circ f(x)) \chi(x) dx, \text{ car } \text{supp } \chi \subset X_0.\end{aligned}$$

Or,  $\phi|_{f(X_0)} \equiv 1$ , donc

$$\begin{aligned}\langle f^*(\phi v), \chi \rangle &= \int_{\mathbb{R}^m} v \circ f(x) \chi(x) dx \\ &= \langle f^*v, \chi \rangle\end{aligned}$$

Et il suit que

$$\langle f^*v, \chi \rangle = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{\phi v}(\eta) I_\chi(\eta) d\eta. \quad (5.19)$$

En plus du lemme de la phase non stationnaire, on va utiliser les deux résultats suivants. Posons  $g_\eta(x) = \langle f(x), \eta \rangle$ . Alors,  $\forall h \in \mathbb{R}^m$ ,

$$\begin{aligned} dg_\eta(x)h &= \left. \frac{d}{dt} g_\eta(x + th) \right|_{t=0} \\ &= \left. \frac{d}{dt} \langle f(x + th), \eta \rangle \right|_{t=0} \\ &= \langle f'(x)h, \eta \rangle \end{aligned}$$

D'où

$$dg_\eta(x)h = \langle h, {}^t f'(x)\eta \rangle \quad (5.20)$$

où  $f'(x)$  désigne la différentielle de  $f$  en  $x$ .

L'ensemble  $\text{supp } \chi$  est un compact inclus dans  $X_0$ , et  $f$  est  $C^\infty$ , donc  $\|{}^t f'(x)\|$  bornée sur  $\text{supp } \chi$  et la borne inférieure  $m$  est atteinte. Comme  $V$  est un cône fermé de  $\mathbb{R}^n \setminus 0$  et pour tout  $x \in X_0$  et  $\forall \eta \in V$ ,  ${}^t f'(x)\eta \neq 0$ , on a que, pour  $x \in \text{supp } \chi$

$$\|{}^t f'(x)\| = \sup_{s \neq 0} \frac{|{}^t f'(x)s|}{|s|} \geq m > 0.$$

En particulier, pour  $\eta \in V$ ,

$$|{}^t f'(x)\eta| \geq m|\eta|.$$

Donc il existe  $C > 0$  une constante telle que

$$|\eta| \leq C|{}^t f'(x)\eta|, \quad \forall x \in \text{supp } \chi, \quad \eta \in V. \quad (5.21)$$

On a ensuite

$$\forall N \in \mathbb{N}^*, \exists C_{N,\chi} > 0 : |I_\chi(\eta)| \leq C_{N,\chi}(1 + |\eta|)^{-N}, \quad \forall \eta \in V. \quad (5.22)$$

Ceci est une conséquence du lemme 1 de la phase non stationnaire. En effet,  $V \subset \mathbb{R}^n \setminus 0$ , donc  $\eta$  ne s'annule pas et on a

$$\begin{aligned} I_\chi(\eta) &= \int_{\mathbb{R}^m} \chi(x) e^{i\langle f(x), \eta \rangle} dx \\ &= \int_{\mathbb{R}^m} \chi(x) e^{i|\eta| \langle f(x), \frac{\eta}{|\eta|} \rangle} dx. \end{aligned}$$

Pour tout  $k \geq 1$ , on applique ce lemme à  $\chi \in C_0^k(X_0)$ ,

$$\phi_\eta : x \mapsto \left\langle f(x), \frac{\eta}{|\eta|} \right\rangle \in C^{k+1}(X_0),$$

qui est à valeurs réelles de gradient non nul (car  $x \in X_0$ ,  $\eta \in V$ ) et à  $\omega = |\eta|$ . Alors il existe une constante  $C_k > 0$  telle que,  $\forall \eta \in V$ ,

$$\begin{aligned} |I_\chi(\eta)| &= \left| \int_{\mathbb{R}^m} \chi(x) e^{i|\eta| \langle f(x), \frac{\eta}{|\eta|} \rangle} dx \right| \\ &\stackrel{(1)}{\leq} C_k \frac{1}{|\eta|^k} \sum_{|\alpha| \leq k} \sup_{X_0} |D^\alpha \chi| \sup_{X_0} |\phi'_\eta(x)|^{|\alpha|-2k}, \end{aligned}$$

avec  $C_k$  indépendante de  $\eta$  car les dérivées successives de  $\phi_\eta$  sont bornées en  $\eta$ .

Il suit

$$\begin{aligned}
|I_\chi(\eta)| &\leq C_k C_{k,\chi} \frac{1}{|\eta|^k} \sum_{|\alpha| \leq k} \sup_{X_0} |\phi'_\eta(x)|^{|\alpha|-2k}, \quad \text{car } \chi \text{ est } C_0^\infty(X_0) \\
&\stackrel{(5.20)}{\leq} C_k C_{k,\chi} \frac{1}{|\eta|^k} \sum_{|\alpha| \leq k} \sup_{X_0} \left| \frac{{}^t f'(x)\eta}{|\eta|} \right|^{|\alpha|-2k} \\
&\leq C_k C_{k,\chi} \frac{1}{|\eta|^k} \sum_{|\alpha| \leq k} \sup_{X_0} \left( \frac{|\eta|}{|{}^t f'(x)\eta|} \right)^{2k-|\alpha|}, \quad \text{avec } 2k - |\alpha| > 0.
\end{aligned}$$

Or,  $\forall x \in \text{supp } \chi, \eta \in V$ ,

$$\frac{|\eta|}{|{}^t f'(x)\eta|} \stackrel{(5.21)}{\leq} C,$$

donc

$$\left( \frac{|\eta|}{|{}^t f'(x)\eta|} \right)^{2k-|\alpha|} \leq C^{2k-|\alpha|}, \quad \forall |\alpha| \leq k.$$

De plus, comme  $V$  est un fermé de  $\mathbb{R}^n \setminus \{0\}$ , le rapport

$$\frac{1 + |\eta|}{|\eta|}$$

reste borné, pour tout  $\eta \in V$ . On en déduit que pour tout  $k$ , il existe une nouvelle constante  $C_{k,\chi}$  qui ne dépend que de  $k$  et de  $\chi$  telle que :

$$|I_\chi(\eta)| \leq C_{k,\chi} (1 + |\eta|)^{-k}, \quad \forall \eta \in V$$

ce qui montre bien (5.22).

Si  $u_j \in C^\infty(Y)$  et  $u_j \rightarrow u$  dans  $\mathcal{D}'_\Gamma(Y)$ , on a,  $\forall N \in \mathbb{N}^*$

$$|\widehat{\phi u_j}| \leq C'_N (1 + |\eta|)^{-N}, \quad \forall \eta \notin V. \quad (5.23)$$

Ceci découle simplement de la définition de la convergence au sens  $\mathcal{D}'_\Gamma(Y)$ . On a enfin qu'il existe  $M$  et une constante  $C > 0$  tels que

$$|\widehat{\phi u_j}(\eta)| \leq C (1 + |\eta|)^M, \quad \forall \eta \in \mathbb{R}^n. \quad (5.24)$$

Ceci découle du principe de la borne uniforme (conséquence de Banach-Steinhaus), qui nous dit que puisque  $u_j \rightarrow u$  dans  $\mathcal{D}'(Y)$ ,  $\forall K$  compact de  $Y$ ,  $\exists C > 0$ ,  $\exists M \in \mathbb{N}$  tels que

$$\forall \psi \in \mathcal{D}(K), \forall j, \quad |\langle u_j, \psi \rangle| \leq C \sup_{|\alpha| \leq M} |\partial^\alpha \psi|. \quad (5.25)$$

Comme  $\phi u_j$  est à support compact, on a que

$$\widehat{\phi u_j}(\eta) = \langle u_j, \phi e^{-i\langle \cdot, \eta \rangle} \rangle, \quad \forall \eta \in \mathbb{R}^n.$$

Ainsi, on utilise (5.25) avec  $\psi = \phi e^{-i\langle \cdot, \eta \rangle}$ , et sachant que

$$\sup_{|\alpha| \leq M} |\partial^\alpha \psi| \leq cte (1 + |\eta|)^{|\alpha|} \leq cte (1 + |\eta|)^M,$$

on obtient (5.24).

Soit donc  $u_j \in C^\infty(Y)$  telle que  $u_j \rightarrow u$  dans  $\mathcal{D}'_\Gamma(Y)$ .

Ainsi on a

$$\begin{aligned}
\langle f^* u_j, \chi \rangle &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{\phi u_j}(\eta) I_\chi(\eta) d\eta \\
&= \frac{1}{(2\pi)^n} \left( \int_V \widehat{\phi u_j}(\eta) I_\chi(\eta) d\eta + \int_{V^c} \widehat{\phi u_j}(\eta) I_\chi(\eta) d\eta \right).
\end{aligned}$$

Comme  $u_j \rightarrow u$ , on a que

$$\langle u_j, \phi e^{-i\langle \cdot, \eta \rangle} \rangle \rightarrow \langle u, \phi e^{-i\langle \cdot, \eta \rangle} \rangle, \quad \forall \eta \in \mathbb{R}^n$$

et

$$\widehat{\phi u_j}(\eta) \rightarrow \widehat{\phi u}(\eta), \quad \forall \eta \in \mathbb{R}^n.$$

De plus,  $\forall j \in \mathbb{N}$ ,  $\phi u_j \in C_0^\infty(Y)$  et les majorations (5.22), (5.23) et (5.24) nous fournissent une domination.

— Dans  $V$  :

$$\forall j \in \mathbb{N}, \eta \in V$$

$$|\widehat{\phi u_j}(\eta)| |I_\chi(\eta)| \leq C_{N,\chi} C(1 + |\eta|)^M (1 + |\eta|)^{-N}, \quad \forall N \geq 1$$

et  $C_{N,\chi} C$  indépendant de  $\eta$ . Pour  $N$  suffisamment grand, on a une domination intégrable de  $|\widehat{\phi u_j}(\eta)| |I_\chi(\eta)|$  et indépendante de  $j$ .

— Dans  $V^c$  :

$$\forall j \in \mathbb{N}, \eta \in V^c$$

$$|\widehat{\phi u_j}(\eta)| |I_\chi(\eta)| \leq C'_N \left( \int_{X_0} |\chi(x)| dx \right) (1 + |\eta|)^{-N}, \quad \forall N \geq 1$$

et  $C'_N \left( \int_{X_0} |\chi(x)| dx \right)$  indépendant de  $\eta$ . Pour  $N$  suffisamment grand, on a une domination intégrable de  $|\widehat{\phi u_j}(\eta)| |I_\chi(\eta)|$  et indépendante de  $j$ .

Par convergence dominée, on conclut que :

$$\langle f^* u_j, \chi \rangle \xrightarrow{j \rightarrow \infty} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{\phi u}(\eta) I_\chi(\eta) d\eta.$$

On a ainsi que  $f^* u_j$  converge dans  $\mathcal{D}'(X_0)$  et on va se ramener au cas général grâce à une partition de l'unité. En effet, si  $\chi \in C_0^\infty(X)$ , on peut écrire

$$\text{supp } \chi = \bigcup_{x_0 \in \text{supp } \chi} \mathring{X}_0$$

où  $X_0$  est un voisinage compact de  $x_0$  d'intérieur non vide. De ce recouvrement, on peut extraire un sous-recouvrement fini par la propriété de Borel-Lebesgue :

$$\text{supp } \chi = \bigcup_{i=1, \dots, d} \mathring{X}_i.$$

Par partition de l'unité, pour tout  $i = 1, \dots, d$ , il existe  $\chi_i \in C_0^\infty(X_i)$  tels que

$$\chi = \sum_{i=1, \dots, d} \chi_i$$

et tels que la somme ne dépende pas de la partition choisie.

Il suit

$$\begin{aligned} \langle f^* u_j, \chi \rangle &= \langle f^* u_j, \sum_{i=1, \dots, d} \chi_i \rangle \\ &= \sum_{i=1, \dots, d} \langle f^* u_j, \chi_i \rangle \end{aligned}$$

et chaque  $\langle f^* u_j, \chi_i \rangle$  converge par ce qui précède.



En définitive,  $\langle f^*u_j, \chi \rangle$  converge dans  $\mathcal{D}'(X)$  vers une limite indépendante de la partition de l'unité choisie. On note  $f^*u$  cette limite.

Montrons maintenant la continuité de l'application

$$f^* : \mathcal{D}'_\Gamma(Y) \rightarrow \mathcal{D}'_{f^*\Gamma}(Y)$$

pour les topologies précédemment définies.

Avec les notations précédentes, on considère  $u_j \rightarrow u$  dans  $\mathcal{D}'_\Gamma(Y)$ ,  $\chi \in C_0^\infty X_0$  et on pose

$$v_j \stackrel{\text{def}}{=} \chi f^*u_j.$$

Alors (5.19) donne, avec  $\chi$  remplacé par  $\chi e^{-i\langle \cdot, \xi \rangle}$ , pour tout  $\xi \in \mathbb{R}^m$  :

$$\begin{aligned} \langle f^*u_j, \chi e^{-i\langle \cdot, \xi \rangle} \rangle &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{\phi u_j}(\eta) I_\chi(\eta, \xi) d\eta \\ I_\chi(\eta, \xi) &= \int_{\mathbb{R}^m} \chi(x) e^{i(\langle f(x), \eta \rangle - \langle x, \xi \rangle)} dx. \end{aligned}$$

D'où, puisque  $v_j$  est à support compact

$$\widehat{v_j}(\xi) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{\phi u_j}(\eta) I_\chi(\eta, \xi) d\eta, \quad \forall \xi \in \mathbb{R}^m.$$

Soit  $W$  un voisinage conique ouvert de  ${}^t f'(x_0)\Gamma_{y_0} = \{ {}^t f'(x_0)\eta \in \mathbb{R}^m \setminus 0 ; (f(x_0), \eta) \in \Gamma \}$ . On peut supposer que  $V$  et  $X_0$  sont tels que :

$${}^t f'(x)\eta \in W, \quad \forall x \in X_0 \quad \text{et} \quad \eta \in V.$$

Puisque  $V$  un voisinage de  $\Gamma_{y_0}$ , on a alors que, pour un  $\epsilon$  suffisamment petit,

$$|{}^t f'(x)\eta - \xi| \geq \epsilon(|\xi| + |\eta|), \quad \forall \xi \notin W, \quad \eta \in V, \quad \text{et} \quad x \in X_0. \quad (5.26)$$

Ici aussi, le lemme 1 de la phase non stationnaire nous donne des majorations :

$$\forall N^* \in \mathbb{N}, \quad \exists C_N > 0,$$

$$|I_\chi(\eta, \xi)| \leq C_N (1 + |\xi| + |\eta|)^{-N}, \quad \forall \xi \notin W, \quad \eta \in V. \quad (5.27)$$

En effet, pour tout  $k \geq 1$ , on applique ce lemme à  $\chi \in C_0^k(X_0)$ ,

$$\phi_{\eta, \xi} : x \mapsto \langle f(x), \frac{\eta}{|\eta| + |\xi|} \rangle - \langle x, \frac{\xi}{|\eta| + |\xi|} \rangle \in C^{k+1}(X_0)$$

qui est à valeurs réelles de gradient non nul, et à  $\omega = |\eta| + |\xi| \neq 0$ . Alors il existe une constante  $C_k$  ne dépendant que de  $k$  telle que,  $\forall \eta \in V, \xi \notin W$  :

$$\begin{aligned} |I_\chi(\eta, \xi)| &= \left| \int_{\mathbb{R}^m} \chi(x) e^{i(|\eta| + |\xi|)(\langle f(x), \frac{\eta}{|\eta| + |\xi|} \rangle - \langle x, \frac{\xi}{|\eta| + |\xi|} \rangle)} dx \right| \\ &\stackrel{(1)}{\leq} C_k \frac{1}{(|\eta| + |\xi|)^k} \sum_{|\alpha| \leq k} \sup_{X_0} |D^\alpha \chi| \sup_{X_0} |\phi'_{\eta, \xi}(x)|^{|\alpha| - 2k}. \end{aligned}$$

Il suit

$$\begin{aligned} |I_\chi(\eta, \xi)| &\leq C_{k, \chi} \frac{1}{(|\eta| + |\xi|)^k} \sum_{|\alpha| \leq k} \sup_{X_0} |\phi'_{\eta, \xi}(x)|^{|\alpha| - 2k}, \quad \text{car } \chi \text{ est } C_0^\infty(X_0) \\ &\stackrel{(5.20)}{\leq} C_{k, \chi} \frac{1}{(|\eta| + |\xi|)^k} \sum_{|\alpha| \leq k} \sup_{X_0} \left( \frac{1}{|\eta| + |\xi|} |{}^t f'(x)\eta - \xi| \right)^{|\alpha| - 2k} \\ &\leq C_{k, \chi} \frac{1}{(|\eta| + |\xi|)^k} \sum_{|\alpha| \leq k} \sup_{X_0} \left( \frac{|\eta| + |\xi|}{|{}^t f'(x)\eta - \xi|} \right)^{2k - |\alpha|}, \quad \text{avec } 2k - |\alpha| > 0. \end{aligned}$$

Or,  $\forall x \in X_0, \eta \in V$  et  $\xi \notin W$

$$\frac{|\eta| + |\xi|}{|{}^t f'(x)\eta - \xi|} \stackrel{(5.26)}{\leq} \frac{1}{\epsilon},$$

donc

$$\left( \frac{|\eta| + |\xi|}{|{}^t f'(x)\eta - \xi|} \right)^{2k-|\alpha|} \leq \frac{1}{\epsilon^{2k-|\alpha|}}, \quad \forall |\alpha| \leq k.$$

De plus, comme  $|\eta| + |\xi| \neq 0$ , le rapport

$$\frac{1 + |\eta| + |\xi|}{|\eta| + |\xi|}$$

reste borné, pour tout  $\eta \in V$  et  $\xi \notin W$ . On en déduit que pour tout  $k$ , il existe une nouvelle constante  $C_{k,\chi}$  qui ne dépend que de  $k$  et de  $\chi$  telle que :

$$|I_\chi(\eta, \xi)| \leq C_N(1 + |\xi| + |\eta|)^{-k}, \quad \forall \xi \notin W, \eta \in V.$$

Ce même lemme nous donne également :

$$\forall N \in \mathbb{N}^*, \exists C'_N > 0,$$

$$|I_\chi(\eta, \xi)| \leq C'_N(1 + |\eta|)^N(1 + |\xi|)^{-N}, \quad \forall \xi \notin W, \eta \notin V. \quad (5.28)$$

En effet, pour tout  $k \geq 1$ , on l'applique à  $\chi e^{i\langle f(x), \eta \rangle} \in C_0^k(X_0)$ ,

$$\phi_\xi : x \mapsto -\langle x, \frac{\xi}{|\xi|} \rangle \in C^{k+1}(X_0)$$

qui est à valeurs réelles de gradient non nul, et à  $\omega = |\xi| \neq 0$ . Alors il existe une constante  $C_k$  ne dépendant que de  $k$  telle que,  $\forall \eta \notin V, \xi \notin W$  :

$$\begin{aligned} |I_\chi(\eta, \xi)| &= \left| \int_{\mathbb{R}^m} \chi(x) e^{i\langle f(x), \eta \rangle} e^{-i|\xi| \langle x, \frac{\xi}{|\xi|} \rangle} dx \right| \\ &\stackrel{(1)}{\leq} C_k \frac{1}{|\xi|^k} \sum_{|\alpha| \leq k} \sup_{X_0} |D^\alpha (\chi e^{i\langle f(x), \eta \rangle})| \sup_{X_0} |\phi'_\xi(x)|^{|\alpha|-2k}. \end{aligned}$$

Il suit

$$|I_\chi(\eta, \xi)| \leq C_{k,\chi} \frac{1}{|\xi|^k} \sum_{|\alpha| \leq k} (1 + |\eta|)^{|\alpha|} \sup_{X_0} \left| \frac{\xi}{|\xi|} \right|^{|\alpha|-2k}, \quad \text{car } \chi \text{ est } C_0^\infty(X_0).$$

De plus, comme  $|\xi| \neq 0$ , le rapport

$$\frac{1 + |\xi|}{|\xi|}$$

reste borné, pour tout  $\xi \notin W$ . On en déduit que pour tout  $k$ , il existe une nouvelle constante  $C_{k,\chi}$  qui ne dépend que de  $k$  et de  $\chi$  telle que :

$$|I_\chi(\eta, \xi)| \leq C_{k,\chi}(1 + |\eta|)^k(1 + |\xi|)^{-k}, \quad \forall \xi \notin W, \eta \notin V. \quad (5.29)$$

On a ainsi,  $\forall N \in \mathbb{N}, \forall \xi \notin W$  :

$$\begin{aligned} |\widehat{v}_j(\xi)| &= \frac{1}{(2\pi)^n} \left| \int_{\mathbb{R}^n} \widehat{\phi u_j}(\eta) I_\chi(\eta, \xi) d\eta \right| \\ &\leq \frac{1}{(2\pi)^n} \left( \left| \int_V \widehat{\phi u_j}(\eta) I_\chi(\eta, \xi) d\eta \right| + \left| \int_{V^c} \widehat{\phi u_j}(\eta) I_\chi(\eta, \xi) d\eta \right| \right) \\ &\leq \frac{1}{(2\pi)^n} \left( \int_V |\widehat{\phi u_j}(\eta)| |I_\chi(\eta, \xi)| d\eta + \int_{V^c} |\widehat{\phi u_j}(\eta)| |I_\chi(\eta, \xi)| d\eta \right) \\ &\stackrel{(5.27)+(5.28)}{\leq} \frac{C_N}{(2\pi)^n} \left( \int_V |\widehat{\phi u_j}(\eta)| (1 + |\xi| + |\eta|)^{-N} d\eta + (1 + |\xi|)^{-N} \int_{V^c} |\widehat{\phi u_j}(\eta)| (1 + |\eta|)^N d\eta \right) \\ &\stackrel{(5.24)}{\leq} \frac{C'_N}{(2\pi)^n} \left( \int_V (1 + |\xi| + |\eta|)^{M-N} d\eta + (1 + |\xi|)^{-N} \int_{V^c} |\widehat{\phi u_j}(\eta)| (1 + |\eta|)^N d\eta \right). \end{aligned}$$

On fixe maintenant  $N \in \mathbb{N}^*$ . Par ce qui précède,  $\exists M \in \mathbb{N}^*$ ,  $\forall N' \in \mathbb{N}^*$ ,  $\exists C'_N > 0$  tels que  $\forall \xi \notin W$ ,

$$\begin{aligned} |\xi|^N |\widehat{v_j}(\xi)| &= |\xi|^N |\widehat{\chi f^* u_j}(\xi)| \\ &\leq C'_{N'} \left( |\xi|^N \int_V (1 + |\xi| + |\eta|)^{M-N'} d\eta \right) \\ &\quad + C'_{N'} \left( (1 + |\xi|)^{-N'} |\xi|^N \int_{V^c} |\widehat{\phi u_j}(\eta)| (1 + |\eta|)^{N'} d\eta \right) \end{aligned}$$

et le membre de droite ne dépend pas de  $j$ .

— Dans  $V$  :

$$\begin{aligned} |\xi|^N \int_V (1 + |\xi| + |\eta|)^{M-N'} d\eta &\leq \int_V (1 + |\xi| + |\eta|)^N (1 + |\xi| + |\eta|)^{M-N'} d\eta \\ &\leq \int_V (1 + |\xi| + |\eta|)^{M+N-N'} d\eta. \end{aligned}$$

On choisit  $N'$  de sorte que  $N' - N - M > n$  et on obtient que

$$\forall \eta \in V, \frac{1}{(1 + |\eta| + |\xi|)^{N'-N-M}} \xrightarrow{|\xi| \rightarrow \infty} 0.$$

De plus,  $\forall \xi \notin W$ ,  $\eta \mapsto \frac{1}{(1 + |\eta| + |\xi|)^{N'-N-M}}$  est continue et

$$\forall \xi \notin W, \eta \in V, \frac{1}{(1 + |\eta| + |\xi|)^{N'-N-M}} \leq \frac{1}{(1 + |\eta|)^{N'-N-M}}$$

qui est intégrable, on peut donc appliquer le théorème de convergence dominée pour obtenir

$$\int_V (1 + |\xi| + |\eta|)^{M+N-N'} d\eta \xrightarrow{|\xi| \rightarrow \infty} 0.$$

— Dans  $V^c$  :

Puisque  $u_j \rightarrow u$  dans  $\mathcal{D}'_T(Y)$ , on a par (5.23)

$$\forall \tilde{N} \in \mathbb{N}, \exists C_{\tilde{N}}, \forall \eta \notin V$$

$$|\widehat{\phi u_j}(\eta)| \leq C_{\tilde{N}} (1 + |\eta|)^{-\tilde{N}}.$$

Alors

$$\begin{aligned} (1 + |\xi|)^{-N'} |\xi|^N \int_{V^c} |\widehat{\phi u_j}(\eta)| (1 + |\eta|)^{N'} d\eta \\ \leq C_{\tilde{N}} (1 + |\xi|)^{-N'} |\xi|^N \int_{V^c} (1 + |\eta|)^{-\tilde{N}} (1 + |\eta|)^{N'} d\eta. \end{aligned}$$

Pour  $N'$  suffisamment grand,

$$(1 + |\xi|)^{-N'} |\xi|^N \xrightarrow{|\xi| \rightarrow 0} 0,$$

et pour  $\tilde{N}$  encore plus grand,

$$\int_{V^c} (1 + |\eta|)^{-\tilde{N}} (1 + |\eta|)^{N'} d\eta < \infty.$$

En sommant les deux cas précédents, on obtient que

$$\sup_{j \in \mathbb{N}} \sup_{\xi \notin W} \widehat{\chi f^* u_j}(\xi) < \infty, \quad \forall N \in \mathbb{N}^*.$$

On a, par définition  $f^* u_j \xrightarrow{\mathcal{D}'(X)} f^* u$  et on a montré

$$\sup_{j \in \mathbb{N}} \sup_{\xi \notin W} \widehat{\chi f^* u_j}(\xi) < \infty, \quad \forall N \in \mathbb{N}^*$$

où  $W^c$  est un cône fermé de  $\mathbb{R}^m$  avec, puisque  $W$  est un voisinage conique de  ${}^t f'(x_0)\Gamma_{y_0}$  et quitte à rétrécir  $\text{supp } \chi$ ,

$$f^*\Gamma \cap (\text{supp } \chi \times W^c) = \emptyset.$$

En se ramenant comme précédemment à une partition de l'unité, on a bien que

$$f^*u_j \longrightarrow f^*u, \text{ dans } \mathcal{D}'_{f^*\Gamma}(X).$$

ce qui montre la continuité de  $f^*$  et achève la preuve.  $\square$

### Produit de distributions

Comme cela a été remarqué dans la partie 5.4.1, nous allons nous appuyer sur la notion de produit tensoriel pour définir le produit de distributions. Il est donc naturel de s'intéresser au front d'onde du produit tensoriel.

#### Théorème 5.4.3 -

Soient  $X$  un ouvert de  $\mathbb{R}^m$  et  $Y$  un ouvert de  $\mathbb{R}^n$ . Si  $u \in \mathcal{D}'(X)$  et  $v \in \mathcal{D}'(Y)$ , on a :

$$WF(u \otimes v) \subset \{(x, y, \xi, \eta) \in X \times Y \times (\mathbb{R}^m \setminus 0) \times (\mathbb{R}^n \setminus 0) ; (x, \xi) \in WF(u) \text{ et } (y, \eta) \in WF(v)\}$$

$$\cup \{(x, y, \xi, 0) \in X \times Y \times (\mathbb{R}^m \setminus 0) \times \{0\} ; (x, \xi) \in WF(u) \text{ et } y \in \text{supp } (v)\}$$

$$\cup \{(x, y, 0, \eta) \in X \times Y \times \{0\} \times (\mathbb{R}^n \setminus 0) ; x \in \text{supp } (u) \text{ et } (y, \eta) \in WF(v)\}.$$

*Démonstration.* D'après la remarque 5.2.3 de la partie précédente, pour vérifier qu'un point  $(y, \eta)$  n'est pas dans le front d'onde de  $u \otimes v$ , il suffit de vérifier que la transformée de Fourier de  $\psi u \otimes v$  est à décroissance rapide dans un cône contenant  $\eta$  pour une seule fonction  $\psi \in C_0^\infty$  telle que  $\psi(y) \neq 0$ . Soient donc  $(x_0, \xi_0) \in X \times (\mathbb{R}^m \setminus 0)$  tel que  $(x_0, \xi_0) \notin WF(u)$  et  $(y_0, \eta_0) \in Y \times (\mathbb{R}^n \setminus 0)$ . Soient aussi  $\varphi_1 \in C_0^\infty(X)$ ,  $\varphi_1(x_0) \neq 0$  à support assez petit et  $\varphi_2 \in C_0^\infty(Y)$ ,  $\varphi_2(y_0) \neq 0$  à support assez petit. Notons que  $\varphi_1 \varphi_2(x_0, y_0) \neq 0$ . On pose :

$$\omega = (\varphi_1 u) \otimes (\varphi_2 v) = [\varphi_{1,x} \varphi_{2,y}] u_x \otimes v_y.$$

Alors  $\omega$  est à support compact et :

$$\begin{aligned} \widehat{\omega}(\xi_1, \xi_2) &= \langle (\varphi_1 u) \otimes (\varphi_2 v), e^{-i(x\xi_1 + y\xi_2)} \rangle \\ &= \langle \varphi_1 u, \langle \varphi_2 v, e^{-i(x\xi_1 + y\xi_2)} \rangle \rangle \\ &= \langle u, \varphi_1 \langle v, \varphi_2 e^{-i(x\xi_1 + y\xi_2)} \rangle \rangle \\ &= \langle u, \varphi_1 \langle v, \varphi_2 e^{-iy\xi_2} \rangle e^{-ix\xi_1} \rangle \\ &= \langle u \varphi_1, e^{-ix\xi_1} \rangle \cdot \langle v \varphi_2, e^{-iy\xi_2} \rangle \\ &= \widehat{\varphi_1 u_1}(\xi_1) \cdot \widehat{\varphi_2 u_2}(\xi_2). \end{aligned}$$

Puisque  $(x_0, \xi_0)$  n'est pas dans le front d'onde de  $u$  et quitte à rétrécir le support de  $\varphi_1$ , il existe  $\Gamma_1 \subset \mathbb{R}^m$  voisinage conique de  $\xi_0$  tel que  $\widehat{\varphi_1 u_1}$  à décroissance rapide sur  $\Gamma_1$ . Puisque  $\varphi_2 v$  à support compact, le théorème 3 nous indique que  $\widehat{\varphi_2 v}$  est à croissance polynomiale, notamment sur un cône ouvert  $\Gamma_2 \in \mathbb{R}^n$  contenant  $\eta_0$ . Alors  $\Gamma_1 \times \Gamma_2$  est un cône de  $\mathbb{R}^m \times \mathbb{R}^n$  contenant  $(\xi_0, \eta_0)$ . Ainsi, il existe  $M \in \mathbb{N}$ ,  $C > 0$  une constante, tels que  $\forall N \in \mathbb{N}$ ,  $\forall (\xi, \eta) \in \Gamma_1 \times \Gamma_2$  :

$$\begin{aligned} |\widehat{\omega}(\xi_1, \xi_2)| &= |\widehat{\varphi_1 u_1}(\xi_1)| |\widehat{\varphi_2 u_2}(\xi_2)| \\ &\leq C(1 + |\xi|)^{-N} (1 + |\eta|)^M. \end{aligned}$$

Les rapports des trois quantités  $|\xi|$ ,  $|\eta|$  et  $|(\xi, \eta)|$  étant bornés dans  $\Gamma_1 \times \Gamma_2$ , on obtient :

$$\forall N \in \mathbb{N}, \exists \tilde{C}, \forall (\xi, \eta) \in \Gamma_1 \times \Gamma_2 : |\widehat{\omega}(\xi_1, \xi_2)| \leq \tilde{C}(1 + |(\xi, \eta)|)^{-N}$$

Ainsi donc  $(x_0, y_0, \xi_0, \eta_0) \notin WF(u \otimes v)$ .

Symétriquement, si  $(x_0, \xi_0) \in X \times (\mathbb{R}^m \setminus 0)$  et  $(y_0, \eta_0) \in Y \times (\mathbb{R}^n \setminus 0)$  tel que  $(y_0, \eta_0) \notin WF(v)$ ,  $(x_0, y_0, \xi_0, \eta_0) \notin WF(u \otimes v)$ . Puisque le front d'onde se projette sur le support singulier, lui-même inclus dans le support, on a :

$$y \notin \text{supp } v \implies (x, y) \notin \text{supp } u \otimes v \implies (x, y, \xi, 0) \notin WF(u \otimes v).$$

$$x \notin \text{supp } u \implies (x, y) \notin \text{supp } u \otimes v \implies (x, y, 0, \eta) \notin WF(u \otimes v).$$

Ceci achève la preuve.  $\square$

Nous avons à présent tous les outils en main pour définir le produit de distributions.

Définition 5.4.7 -

Soit un sous-ensemble  $M \in \mathbb{R}^m \times (\mathbb{R}^m \setminus 0)$ . On définit l'*antipodal*  $M^a$  de  $M$  comme :

$$M^a \stackrel{\text{def}}{=} \{(x, -\xi) ; (x, \xi) \in M\}.$$

Théorème 5.4.4 -

Soit  $X$  un ouvert de  $\mathbb{R}^m$ . Soient  $u, v \in \mathcal{D}'(X)$ . Alors le produit  $uv$  est bien défini sous l'hypothèse suivante :

$$WF(u) \cap (WF(v))^a = \emptyset.$$

Ce produit est défini comme l'image réciproque du produit tensoriel  $u \otimes v$  par la diagonale

$$\begin{aligned} d : X &\longrightarrow X \times X. \\ x &\longmapsto (x, x) \end{aligned}$$

Dans ce cas, on a :

$$WF(uv) \subset \{(x, \xi) \in X \times (\mathbb{R}^m \setminus 0) ; \exists \eta_1, \eta_2 \in \mathbb{R}^m, \xi = \eta_1 + \eta_2,$$

$$[(x, \eta_1) \in WF(u) \text{ ou } \eta_1 = 0] \text{ et } [(x, \eta_2) \in WF(v) \text{ ou } \eta_2 = 0]\}$$

*Démonstration.* Soient  $u$  et  $v$  vérifiant les hypothèses. Alors  $u \otimes v \in \mathcal{D}'(X \times X)$  et

$N_d = \{(x, x, \xi_1, \xi_2) \in X \times X \times \mathbb{R}^m \times \mathbb{R}^m ; {}^t d'(x)(\xi_1, \xi_2) = 0\}$ . Soit  $x \in X$  et  $(\xi_1, \xi_2) \in \mathbb{R}^m \times \mathbb{R}^m$ .

On a :

$$\begin{aligned} (x, x, \xi_1, \xi_2) \in N_d &\iff {}^t d'(x) \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = 0 \\ &\iff \begin{pmatrix} 1 & 0 & 1 & 0 \\ & \ddots & & \ddots \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = 0 \\ &\iff \xi_1 + \xi_2 = 0 \end{aligned}$$

Il suit que  $N_d = \{(x, x, \xi, -\xi) ; x \in X \text{ et } \xi \in \mathbb{R}^m\}$ .

Montrons que  $N_d \cap WF(u \otimes v) = \emptyset$ , ce qui, par le théorème 2, montrera que l'image réciproque de  $u \otimes v$  par  $d$  est définie.

Soit  $(x_1, x_2, \xi_1, \xi_2) \in N_d \cap WF(u \otimes v)$ . Alors, par ce qui précède,  $x_1 = x_2$  et  $\xi_1 = -\xi_2$ . Notons que  $(x_1, x_1, \xi_1, -\xi_1) \in WF(u \otimes v)$ , et que donc  $\xi_1 \neq 0$ . Par le théorème 3,  $(x, \xi_1) \in WF(u)$  et  $(x, -\xi_1) \in WF(v)$ .

Donc  $(x_1, \xi_1) \in WF(u) \cap (WF(v))^a$ , ce qui est absurde par hypothèse. Il suit que le produit, en tant qu'image réciproque, est bien défini.

Fixons maintenant  $(x_0, \xi_0) \in WF(uv) = WF(d^*u \otimes v)$ . Alors, par le corollaire du théorème 2,  $(x_0, \xi_0) \in d^*WF(u \otimes v) = \{(x, {}^t d'(x)\eta) ; (d(x), \eta) \in WF(u \otimes v)\}$ . Ainsi il existe  $x \in X$ , et  $\eta \in \mathbb{R}^{2m} \setminus 0$  tels que  $(d(x), \eta) \in WF(u \otimes v)$  et  $(x_0, \xi_0) = (x, {}^t d'(x)\eta)$ . Donc  $x_0 = x$  et  $\xi_0 = {}^t d'(x)\eta$ . Comme  $\eta \in \mathbb{R}^{2m} \setminus 0$ , on peut noter  $\eta = (\eta_1, \eta_2)$ , avec  $\eta_1, \eta_2 \in \mathbb{R}^m$  non simultanément nuls et les mêmes calculs que précédemment montrent que  $\xi_0 = {}^t d'(x)(\eta_1, \eta_2) = \eta_1 + \eta_2$ . On a ainsi  $(x_0, \xi_0) = (x, \eta_1 + \eta_2)$  et  $(x, \eta_1, \eta_2) \in WF(u \otimes v)$ . Par le théorème 3, trois cas se produisent alors :

$$\begin{aligned} (x, \eta_1) &\in WF(u) \text{ et } (x, \eta_2) \in WF(v) \\ \text{ou } \eta_2 &= 0, (x, \eta_1) \in WF(u) \text{ et } x \in \text{supp } v \\ \text{ou } \eta_1 &= 0, (x, \eta_2) \in WF(v) \text{ et } x \in \text{supp } u, \end{aligned}$$

ce qui montre l'inclusion voulue et achève la preuve.  $\square$

*Remarque.* Puisque les images réciproques ont été définies par une extension continue de composées, cette définition étend la multiplication d'une distribution avec une fonction  $C^\infty$  vue précédemment. De plus, les règles de calcul standard sont toujours vérifiées pour cette définition étendue.

Terminons cette partie par un exemple.

**Exemples 5.4.2.** Calcul du produit de deux mesures de densité linéiques :

$$\begin{aligned} u : \varphi &\mapsto \int_{\mathbb{R}} \varphi(0, y) dy \in \mathcal{D}'(\mathbb{R}^2) \\ v : \varphi &\mapsto \int_{\mathbb{R}} \varphi(x, 0) dx \in \mathcal{D}'(\mathbb{R}^2). \end{aligned}$$

Comme vu dans l'exemple (5.4.1), on a que

$$\begin{aligned} u &= \delta_0 \otimes \mathbb{1}_{\mathbb{R}} \\ v &= \mathbb{1}_{\mathbb{R}} \otimes \delta_0. \end{aligned}$$

Alors, le théorème 3 et l'exemple (5.2.1) nous donnent :

$$\begin{aligned} WF(u) &\subset \{0\} \times \mathbb{R} \times (\mathbb{R} \setminus 0) \times \{0\} \\ WF(v) &\subset \mathbb{R} \times \{0\} \times \{0\} \times (\mathbb{R} \setminus 0) \end{aligned}$$

Si  $(x, y, \xi, \eta) \in WF(u) \cap (WF(v))^a$ , alors

$$\begin{aligned} (x, y, \xi, \eta) &= (0, y, \xi, 0) \\ (x, y, \xi, \eta) &= (x, 0, 0, \eta), \end{aligned}$$

donc  $(x, y, \xi, \eta)$  est nul, ce qui est impossible. On en déduit que

$$WF(u) \cap (WF(v))^a = \emptyset$$

et donc que le produit  $uv$  est bien défini.

Construisons  $u_j \in C_0^\infty(\mathbb{R}^2)$  et  $v_j \in C_0^\infty(\mathbb{R}^2)$  telles que

$$u_j \xrightarrow{\mathcal{D}_{\Gamma_1}} u \text{ et } v_j \xrightarrow{\mathcal{D}_{\Gamma_2}} v.$$

Alors on aura que

$$u_j \otimes v_j \xrightarrow{\mathcal{D}'_{\Gamma}} u \otimes v.$$

On prend donc

$$\begin{aligned} f_j &\longrightarrow \delta_0, \quad f_j \in C_0^\infty(\mathbb{R}) \\ h_j &\longrightarrow \mathbf{1}_{\mathbb{R}}, \quad h_j \in C_0^\infty(\mathbb{R}) \end{aligned}$$

telles que

$$\begin{aligned} u_j &\stackrel{\text{def}}{=} f_j \otimes h_j \xrightarrow{\mathcal{D}'_{\Gamma_1}} u \\ v_j &\stackrel{\text{def}}{=} h_j \otimes f_j \xrightarrow{\mathcal{D}'_{\Gamma_2}} v. \end{aligned}$$

Soit  $\phi \in C_0^\infty(\mathbb{R}^2)$  de la forme

$$\phi(x, y) = \varphi(x)\psi(y),$$

avec  $\varphi, \psi \in C_0^\infty(\mathbb{R})$ .

On a

$$\begin{aligned} \langle d^*(u_j \otimes v_j), \phi \rangle &= \int_{\mathbb{R}^2} u_j(x, y) v_j(x, y) \phi(x, y) dx dy \\ &= \int_{\mathbb{R}^2} f_j(x) h_j(y) h_j(x) f_j(y) \varphi(x) \psi(y) dx dy \\ &= \int_{\mathbb{R}} f_j(x) h_j(x) \varphi(x) dx \int_{\mathbb{R}} h_j(y) f_j(y) \psi(y) dy, \text{ par Fubini.} \end{aligned}$$

Or, quitte à prendre  $j$  grand, on peut supposer que  $h_j|_{\text{supp } f_j} \equiv 1$ . Donc

$$\langle d^*(u_j \otimes v_j), \phi \rangle = \int_{\mathbb{R}} f_j(x) \varphi(x) dx \int_{\mathbb{R}} f_j(y) \psi(y) dy \longrightarrow \varphi(0)\psi(0) = \phi(0, 0).$$

Par densité des sommes  $\sum_k \varphi_k(x) \psi_k(y)$  dans  $C_0^\infty(\mathbb{R}^2)$ , on a montré que

$$\forall \phi \in C_0^\infty(\mathbb{R}^2), \quad \langle d^*(u_j \otimes v_j), \phi \rangle \longrightarrow \delta_{(0,0)}(\phi).$$

Donc on a bien que

$$uv = \delta_{(0,0)}.$$

---

## 5.5. Front d'onde et équations aux dérivées partielles

---

Soit  $\Omega$  un ouvert de  $\mathbb{R}^n$ .

Cadre : On se donne une équation aux dérivées partielles linéaire d'ordre  $m$ , à coefficients  $a_\alpha$  réels et de classe  $C^\infty$

$$P(x, \partial_x)u \stackrel{\text{def}}{=} \sum_{|\alpha| \leq m} a_\alpha(x) \partial^\alpha u(x) \stackrel{\mathcal{D}'(\Omega)}{=} f(x) \quad (5.30)$$

On associe plusieurs invariants à cette équation :

◇ le symbole principal :  $p_m(x, \xi) = \sum_{|\alpha|=m} a_\alpha(x) \xi^\alpha$

C'est une fonction  $C^\infty$  sur  $\Omega \times \mathbb{R}^n$  polynomiale et homogène de degré  $m$  en  $\xi = (\xi_1, \dots, \xi_n)$

◇ la variété caractéristique :

$$\text{Car}(P) = \{(x, \xi) \in \Omega \times (\mathbb{R}^n \setminus \{0\}), p_m(x, \xi) = 0\}$$

### 5.5.1. Théorème fondamental (Hörmander)

Théorème 5.5.1 -

| Si  $u \in \mathcal{D}'(\Omega)$  est solution de (5.30) alors on a  $WF(u) \subset WF(f) \cup \text{Car}(P)$

*Démonstration.* On va raisonner par contraposée i.e. montrer que

$$(WF(f))^c \cap (\text{Car}(P))^c \subset (WF(u))^c.$$

Commençons donc par prendre un point  $(x_0, \xi_0) \in \Omega \times (\mathbb{R}^n \setminus \{0\})$  dans  $(WF(f))^c \cap (\text{Car}(P))^c$ . Alors  $p_m(x_0, \xi_0) \neq 0$ .

Lemme 5.5.1 -

| Il existe une constante positive  $C$ , un voisinage  $U_{x_0} \subset \Omega$  de  $x_0$  et un cône ouvert  $V_{\xi_0}$  contenant  $\xi_0$  tel que :

$$|\xi|^m \leq C |p_m(x, \xi)| \quad \forall x \in U_{x_0}, \quad \forall \xi \in V_{\xi_0} \quad (5.31)$$

▷ En effet, comme  $p_m$  est homogène en  $\xi$ , il suffit de le montrer pour  $\xi \in V_{x_0} \cap \mathbb{S}^1$ .

Or sur  $\mathbb{S}^1$ ,  $|p_m(x, \xi)|$  est minoré et ne s'annule pas pour  $x \in U_{x_0}$  un voisinage de  $x_0$  et  $\xi \in V_{\xi_0}$  un voisinage de  $\xi_0$  donc il existe  $c$  tel que  $|p_m(x, \xi)| \geq c$  ce qui conclut en posant  $C = \frac{1}{c}$ .  $\square$

On veut montrer que  $(x_0, \xi_0) \in (WF(u))^c$  i.e. que  $(x_0, \xi_0)$  est un point où  $u$  est microlocalement de classe  $C^\infty$ . Soit  $\phi \in C_0^\infty(U_{x_0})$  tel que  $\phi(x_0) = 1$ .

On va estimer  $\widehat{\phi u}(\xi)$  pour  $\xi \in V_{\xi_0}$ . On note  ${}^tP$  l'adjoint de  $P$  définie par  $\forall u, v \in \mathcal{D}'(\Omega)$ ,  $\langle Pu, v \rangle := \langle u, {}^tPv \rangle$ .

Comme ici  $Pu = \sum_{|\alpha| \leq m} a_\alpha(x) \partial^\alpha u(x)$ , on a donc que  ${}^tPv = \sum_{|\alpha| \leq m} a_\alpha(x) (-\partial)^\alpha v(x)$ .

Ainsi,

$$Pu = f \iff \langle u, {}^tPv \rangle = \langle f, v \rangle \quad \forall v \in \mathcal{D}'(\Omega)$$

On cherche donc  $v$  tel que  $\langle u, {}^tPv \rangle = \widehat{\phi u}(\xi)$  car on sait estimer le membre de droite. Si  ${}^tPv(x) = \phi(x) e^{-i\langle x, \xi \rangle}$ , alors on a bien que

$$\langle u, {}^tPv \rangle = \int u(x) \phi(x) e^{-i\langle x, \xi \rangle} dx = \widehat{\phi u}(\xi)$$

A première vue, pour  $\xi$  très grand, une solution approchée pour  $v$  semble être

$$e^{-i\langle x, \xi \rangle} \phi(x) / p_m(x, \xi).$$

Néanmoins, essayons de trouver une solution valable pour tout  $\xi$ . On va poser en toute généralité

$$v(x) = w(x) e^{-i\langle x, \xi \rangle} / p_m(x, \xi)$$

où  $w$  est à déterminer.



Lemme 5.5.1 -

Soit  $Q(x, D)$  un opérateur différentiel linéaire. On a alors pour toute fonction  $a$ , pour tout  $(x, \xi) \in \Omega \times (\mathbb{R}^n \setminus \{0\})$  :

$$Q(x, D)(ae^{i\langle x, \xi \rangle}) = e^{i\langle x, \xi \rangle} Q(x, D + \xi)(a) \quad (5.32)$$

▷ Par linéarité de  $Q$ , il suffit de le faire pour les  $D^\alpha$  :

$$\begin{aligned} D^\alpha (ae^{i\langle x, \xi \rangle}) &= \sum_{|\beta| \leq \alpha} \binom{\alpha}{\beta} D^\beta a D^{\alpha-\beta} (e^{i\langle x, \xi \rangle}) \\ &= \sum_{|\beta| \leq \alpha} \binom{\alpha}{\beta} D^\beta a \xi^{\alpha-\beta} (e^{i\langle x, \xi \rangle}) \\ &= e^{i\langle x, \xi \rangle} (D + \xi)^\alpha (a) \quad \square \end{aligned}$$

En appliquant (5.32) avec  $Q = {}^t P$ ,  $a = \frac{w}{p_m(x, \xi)}$ ,  $D = \partial$  et pour  $\xi \leftarrow -\xi$ , on obtient que :

$$\begin{aligned} {}^t P \left( \frac{we^{-i\langle x, \xi \rangle}}{p_m(x, \xi)} \right) &= e^{-i\langle x, \xi \rangle} {}^t P(x, \partial - \xi) \left( \frac{w}{p_m(x, \xi)} \right) \\ &= e^{-i\langle x, \xi \rangle} \sum_{|\alpha| \leq m} (-\partial + \xi)^\alpha \left( \frac{a_\alpha w}{p_m(x, \xi)} \right) \\ &= e^{-i\langle x, \xi \rangle} \sum_{|\alpha| \leq m} \sum_{\beta \leq \alpha} \xi^\beta \binom{\alpha}{\beta} (-\partial_x)^{\alpha-\beta} \left( \frac{a_\alpha w}{p_m(x, \xi)} \right) \\ &= e^{-i\langle x, \xi \rangle} \left( \sum_{|\alpha|=m} a_\alpha \xi^\alpha \left( \frac{w}{p_m(x, \xi)} \right) + \sum_{1 \leq j \leq m} \sum_{|\beta|=m-j} \xi^\beta \sum_{\alpha \geq \beta} \binom{\alpha}{\beta} (-\partial_x)^{\alpha-\beta} \left( \frac{a_\alpha w}{p_m(x, \xi)} \right) \right) \\ &= e^{-i\langle x, \xi \rangle} \left( \sum_{|\alpha|=m} a_\alpha \xi^\alpha \left( \frac{w}{p_m(x, \xi)} \right) + \sum_{1 \leq j \leq m} R_j(w) \right) \end{aligned}$$

en posant

$$R_j(w) = \sum_{|\beta|=m-j} \xi^\beta \sum_{\alpha \geq \beta} \binom{\alpha}{\beta} (-\partial_x)^{\alpha-\beta} \left( \frac{a_\alpha w}{p_m(x, \xi)} \right)$$

où  $R_j$  est homogène en  $\xi$  de degré  $-j$  et  $R_j|\xi|^j$  est un opérateur différentiel d'ordre  $\leq j$  car

$$R_j(w)|\xi|^j = \sum_{|\beta|=m-j} |\xi|^j \xi^\beta \sum_{\alpha \geq \beta, \gamma \leq \alpha-\beta} \binom{\alpha}{\beta} \binom{\alpha-\beta}{\alpha-\beta-\gamma} \partial^{\alpha-\beta-\gamma} \left( \frac{a_\alpha}{p_m(x, \xi)} \right) \partial_x^\gamma w(x)$$

donc  $|\gamma| = |\alpha - \beta| = |\alpha| - |\beta| \leq m - (m - j) = j$ .

On remarque que  $\sum_{|\alpha|=m} a_\alpha \xi^\alpha \left( \frac{w}{p_m(x, \xi)} \right) = w$ , donc comme  ${}^t P v(x) = \phi(x) e^{-i\langle x, \xi \rangle}$ , on obtient pour  $w$  une équation de la forme :

$$w - R w = \phi \quad (5.33)$$

où  $R = -(R_1 + \dots + R_m)$ .

Grâce à (5.31), toutes les dérivées en  $x$  des coefficients de  $R_j|\xi|^j$  sont bornées dans  $U_{x_0} \times V_{\xi_0}$ , puisque  $R_j|\xi|^j$  est homogène en  $\xi$  de degré 0.

Formellement, l'équation (5.33) est satisfaite pour  $w = \sum R^k \phi$ . Cependant, comme la somme ne converge pas a priori, on va prendre une somme partielle à la place. On pose donc

$$w_N = \sum_{k < N} R^k \phi$$

on a alors :

$$w_N - R w_N = \phi - R^N \phi \quad (5.34)$$

où  $R^N$  est une somme de termes qui contiennent tous un facteur  $|\xi|^{-k}$  pour un certain  $k \geq N$ . L'équation précédente signifie que

$${}^t P \left( \frac{w_N e^{-i\langle x, \xi \rangle}}{p_m(x, \xi)} \right) = e^{-i\langle x, \xi \rangle} (\phi - R^N \phi)$$

Donc

$$\langle f, \left( \frac{w_N e^{-i\langle \cdot, \xi \rangle}}{p_m(x, \xi)} \right) \rangle = \langle u, {}^t P \left( \frac{w_N e^{-i\langle \cdot, \xi \rangle}}{p_m(x, \xi)} \right) \rangle = \widehat{\phi u}(\xi) - \langle u, e^{-i\langle \cdot, \xi \rangle} R^N \phi \rangle$$

D'où

$$\widehat{\phi u}(\xi) = \langle u, e^{-i\langle \cdot, \xi \rangle} R^N \phi \rangle + \langle f, \left( \frac{w_N e^{-i\langle \cdot, \xi \rangle}}{p_m(x, \xi)} \right) \rangle, \quad \forall \xi \in V_{\xi_0} \quad (5.35)$$

Si la distribution  $u$  est d'ordre  $\mu$  sur un voisinage du support de  $\phi$ , alors le premier terme du membre de droite de (5.35) peut être contrôlé par

$$C \sum_{|\alpha| \leq \mu} \sup |\partial^\alpha (e^{-i\langle \cdot, \xi \rangle} R^N \phi)| \leq C_N |\xi|^{\mu-N}, \quad |\xi| \geq 1$$

Ici  $N - \mu$  est aussi grand que souhaité.

Par hypothèse,  $(x_0, \xi_0) \in (WF(f))^c$ , donc on peut choisir le voisinage  $U_{x_0}$  de  $x_0$  et le voisinage conique  $V_{\xi_0}$  de  $\xi_0$  tel que pour un certain entier  $M$ , et  $k = 1, 2, \dots$

$$\sup_V |\xi|^k |\widehat{\psi f}(\xi)| \leq C_k \sum_{|\alpha| \leq k+M} \sup |\partial^\alpha \psi|, \quad \forall \psi \in \mathcal{D}'(U_{x_0}) \quad (5.36)$$

En effet, fixons  $\phi \in \mathcal{D}(\Omega)$  tel que  $\phi = 1$  sur un voisinage  $U_{x_0} \subset w$  de  $x_0$ . Soit  $\psi \in \mathcal{D}(U_{x_0})$ . On a donc que  $\psi = \psi \phi$ . Ainsi,

$$\widehat{\psi f}(\xi) = \widehat{\psi \phi f}(\xi) = (2\pi)^{-n} \int \hat{\psi}(\eta) \widehat{\phi f}(\xi - \eta) d\eta.$$

Pour un certain  $M \geq 0$  on a

$$|\widehat{\phi f}(\xi)| \leq C(1 + |\xi|)^M.$$

Soit  $c \in ]0, 1[$ . On va couper l'intégrale en deux en fonction de si  $|\eta| < c|\xi|$  et  $|\eta| \geq c|\xi|$ . On a alors :

$$\begin{aligned} (2\pi)^n |\widehat{\psi f}(\xi)| &\leq \int_{|\eta| < c|\xi|} |\hat{\psi}(\eta) \widehat{\phi f}(\xi - \eta)| d\eta + \int_{|\eta| \geq c|\xi|} |\hat{\psi}(\eta) \widehat{\phi f}(\xi - \eta)| d\eta \\ &\leq \sup_{|\xi - \eta| < c|\xi|} |\widehat{\phi f}| \|\hat{\psi}\|_{L^1} + \int_{|\eta| \geq c|\xi|} |\hat{\psi}(\eta) \widehat{\phi f}(\xi - \eta)| d\eta \quad (\text{changement de variable}) \\ &\leq \sup_{|\xi - \eta| < c|\xi|} |\widehat{\phi f}| \|\hat{\psi}\|_{L^1} + C \int_{|\eta| \geq c|\xi|} |\hat{\psi}(\eta)| (1 + |\eta|)^M (1 + c^{-1})^M d\eta \end{aligned}$$

car si  $|\eta| \geq c|\xi|$ , alors  $|\xi - \eta| \leq (1 + c^{-1})|\eta|$  et donc :

$$\begin{aligned} |\widehat{\phi f}(\xi - \eta)| &\leq C(1 + |\xi - \eta|)^M \\ &\leq C(1 + (1 + c^{-1})|\eta|)^M \\ &\leq C(1 + |\eta|)^M \left(1 + \frac{|\eta|c^{-1}}{1 + |\eta|}\right)^M \\ &\leq C(1 + |\eta|)^M (1 + c^{-1})^M \end{aligned}$$

Si  $\Gamma$  est un cône ouvert dans lequel  $\widehat{\phi f}$  est à décroissance rapide, et  $\Gamma_1$  est un cône fermé inclus dans  $\Gamma \cup \{0\}$  on peut choisir  $c$  tel que  $\eta \in \Gamma$  si  $\xi \in \Gamma_1$  et  $|\xi - \eta| < c|\xi|$ , ce qui est évidemment possible si  $|\xi| = 1$ . On a pour  $N \geq 0$  :

$$\begin{aligned} \sup_{\Gamma_1} (1 + |\xi|)^N |\widehat{\psi f}(\xi)| &\leq (1 - c)^{-N} \sup_{\Gamma} |\widehat{\phi f}(\eta)| (1 + |\eta|)^N \|\widehat{\psi}\|_{L^1} \\ &\quad + C(1 + c^{-1})^{N+M} \int |\widehat{\psi}(\eta)| (1 + |\eta|)^{N+M} d\eta \end{aligned}$$

Or,

$$\forall \alpha, |\eta|^\alpha |\widehat{\psi}(\eta)| = \left| \int e^{-ix \cdot \eta} \partial^\alpha \psi(x) dx \right| \leq |\text{supp}(\psi)| \|\partial^\alpha \psi\|_\infty$$

Donc pour  $|\alpha| = N + M + n + 1$ ,  $|\eta| \geq 1$  :

$$\begin{aligned} \int |\widehat{\psi}(\eta)| (1 + |\eta|)^{N+M} d\eta &= \int |\widehat{\psi}(\eta)| \sum_{|\gamma| \leq N+M} C_\gamma |\eta|^\gamma d\eta \\ &= \int \sum_{|\gamma| \leq N+M} C_\gamma |\widehat{\psi}(\eta)| |\eta|^\gamma d\eta \\ &\leq \int \sum_{|\gamma| \leq N+M} C_\gamma |\widehat{\psi}(\eta)| |\eta|^\alpha |\eta|^{-n-1} d\eta \\ &\leq \sum_{|\gamma| \leq N+M} C_\gamma |\text{supp}(\psi)| \|\partial^\alpha \psi\|_\infty \int |\eta|^{-n-1} d\eta \\ &\leq C |\text{supp}(\psi)| \sup_{|\alpha| \leq N+M+n+1} \|\partial^\alpha \psi\|_\infty \end{aligned}$$

D'autre part,  $(1 - c)^{-N} \sup_{\Gamma} |\widehat{\phi f}(\eta)| (1 + |\eta|)^N \|\widehat{\psi}\|$  étant aussi majoré par  $\sup_{|\alpha| \leq N+M+n+1} \|\partial^\alpha \psi\|_\infty$  on a bien prouvé (5.36).

Prenons  $\psi = \frac{w_N}{p_m(\cdot, \xi)}$ , on conclue alors que le deuxième terme du membre de droite de (5.35) est un  $O(|\xi|^{-k})$  lorsque  $\xi \rightarrow \infty$  dans  $V$ .

Ainsi,

$$\widehat{\phi u} = O(|\xi|^{-k}), \quad \xi \in V, \quad k = 1, 2, \dots$$

On a montré que  $(x_0, \xi_0)$  est un point où  $u$  est microlocalement de classe  $C^\infty$  donc que  $(x_0, \xi_0) \in (WF(u))^c$  ce qui conclue la preuve.  $\square$

### 5.5.2. Cas d'un opérateur elliptique

*Définition 5.5.1 -*

Un opérateur différentiel  $L$  d'ordre  $m$  dans un domaine  $\Omega$  de  $\mathbb{R}^n$  défini par  $Lu = \sum_{|\alpha|=m} a_\alpha(x) \partial^\alpha u$  est dit *elliptique* si pour tout  $x \in \Omega$  et pour tout  $\xi \in \mathbb{R}^n$  non nul, on a

$$\sum_{|\alpha|=m} a_\alpha(x) \xi^\alpha \neq 0.$$

On a donc immédiatement le résultat suivant :

Corollaire 5.5.1 -

Si  $P$  est elliptique, alors

$$WF(u) = WF(Pu) \quad \text{et donc on a aussi que}$$

$$\text{supp sing}(u) = \text{supp sing}(Pu) \quad \text{pour tout } u \in \mathcal{D}'(\Omega).$$

---

## Liste des références

---

Lars Hörmander. *The Analysis of Linear Partial Differential Operators*, chapter I. Springer Berlin Heidelberg, 1983.

# Algorithme DQN



Auriane GABAUT  
Février 2023

## Sommaire

---

|                                                             |            |
|-------------------------------------------------------------|------------|
| <b>Introduction et Motivations</b>                          | <b>105</b> |
| <b>6.1 Préliminaires</b>                                    | <b>107</b> |
| 6.1.1 Bases de <i>Reinforcement Learning</i>                | 107        |
| 6.1.2 <i>Q</i> -Learning                                    | 110        |
| <b>6.2 Deep <i>Q</i>-Learning</b>                           | <b>111</b> |
| 6.2.1 Le Deep <i>Q</i> -Learning et les réseaux de neurones | 112        |
| 6.2.2 Le <i>Q</i> -Network                                  | 113        |
| 6.2.3 Mise en place de l'Algorithme DQN                     | 114        |
| 6.2.4 Double <i>Q</i> -Learning                             | 118        |
| <b>6.3 Algorithme Rainbow</b>                               | <b>122</b> |
| 6.3.1 Prioritized Replay                                    | 122        |
| 6.3.2 Dueling Networks                                      | 122        |
| 6.3.3 Multi-Step Learning                                   | 125        |
| 6.3.4 Distributional RL                                     | 125        |
| 6.3.5 Noisy Nets                                            | 126        |
| 6.3.6 Rainbow                                               | 128        |

---

---

## Introduction et Motivations

---

### Introduction

L'apprentissage par renforcement (ou *Reinforcement Learning*, RL) consiste en un agent qui, à force d'entraînement et donc à force d'acquérir de l'expérience, décide des actions à prendre selon l'état de l'environnement dans lequel il évolue. Une récompense, positive ou négative, lui est alors retournée selon l'action prise et son effet sur l'environnement. L'objectif de l'agent sera alors d'optimiser les décisions qu'il prend pour maximiser la somme de ses récompenses.

Un premier cadre populaire d'apprentissage par renforcement est l'univers des jeux vidéos. L'environnement serait alors le jeu dans lequel l'agent évolue ; les actions qu'il peut prendre sont les différentes possibilités de mouvements dans ce jeu, contraintes par les règles de celui-ci ; les récompenses peuvent être associées à la réussite d'une mission, d'un niveau, ou à un "Game Over", une pénalité, des dégâts, etc [1, 2].

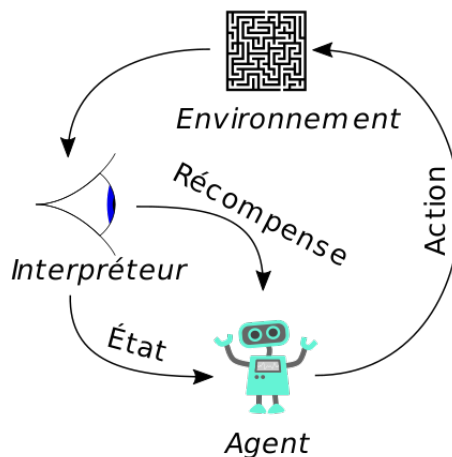


FIGURE 6.1 – Les composants d'un système de R.L. [2].

### Motivations

"Rainbow : Combining Improvements in Deep Reinforcement Learning" [3] est un article publié en 2017 qui présente une amélioration de l'algorithme DQN (Deep *Q*-Network) combinant plusieurs améliorations qui ont été proposées précédemment dans la littérature. Les auteurs ont combiné 7 améliorations différentes, déjà existantes, pour créer l'algorithme "Rainbow". Celui-ci montre des résultats bien meilleurs que l'algorithme original. Les améliorations utilisées sont le Double *Q*-Learning, Prioritized Replay, Dueling Network, Multi-Step Learning, Distributional Reinforcement Learning, et Noisy Nets.

Dans le cadre de l'article, l'algorithme DQN est entraîné et testé sur les jeux Atari. Les jeux Atari sont une série de jeux vidéos développés et commercialisés par Atari, Inc. dans les années 1970 et 1980. Ces jeux ont été très populaires à l'époque et ont contribué à populariser les consoles de salon. Plusieurs de ces jeux, tels que *PacMan*, *Space Invaders* et *Asteroids*, sont devenus des icônes de la culture populaire. Ces jeux sont un cadre simple et très accessible dans lequel faire du Reinforcement Learning. L'article utilise le package `gymnasium` [4] qui implémente complètement les environnements de ces jeux et regroupe un grand nombre de méthodes (e.g. `.n_actions`, `.action_space`, etc.).

Notre objectif ici est de reprendre les idées données par les auteurs de l'article en présentant, en amont, les notions et concepts nécessaires à sa compréhension. Nous avons pu tester deux algorithmes d'apprentissage évoqués dans l'article dont nous comparerons les performances. Faute d'une bonne installation, nous proposons un entraînement et des tests sur un environnement plus simple, inspiré du jeu *Snake*, dont l'implémentation est expliquée en annexe B.

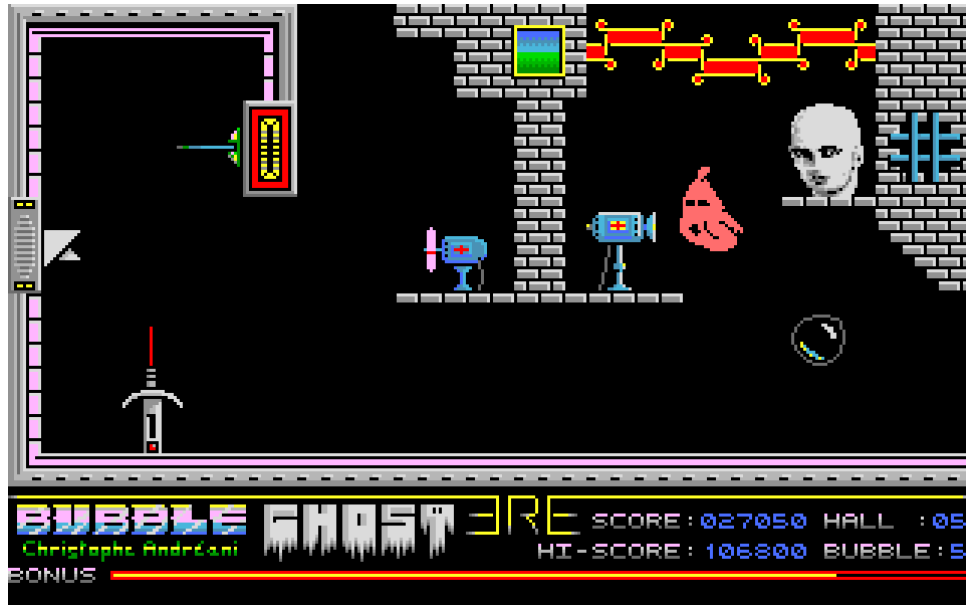


FIGURE 6.2 – Image du tableau n°5 du jeu vidéo *Bubble Ghost* sorti sur Atari ST en 1987.

Auteur du jeu : Christophe Andréani.

### Remerciements

Pour réaliser ce rapport, nous avons d'abord testé les algorithmes implémentés sur nos ordinateurs personnels et nous nous sommes malheureusement vite rendus compte que cela les ralentissait et allait prendre du temps. Il s'est alors offert à nous deux choix : ne faire que peu d'épisodes sur les deux algorithmes présentés, mais conserver nos ordinateurs pour travailler ou l'inverse...

Par chance, et nous les en remercions amplement, Ales BRAHITI (M2 FES) et Clémentine CHAZAL (M2 MVA) nous ont spontanément et volontairement prêtés d'anciens ordinateurs, que nous avons pu laisser jouer plus longtemps.

---

## 6.1. Préliminaires

---

### 6.1.1. Bases de Reinforcement Learning

Nous commençons par rappeler les éléments nécessaires à la compréhension de ce rapport et fixons les notations [1, 2].

## Généralités

L'apprentissage par renforcement ou *Reinforcement Learning* est un domaine de l'apprentissage automatique dont le but est d'entraîner un agent à entreprendre les meilleures actions possibles étant donné une situation. La performance d'une action est renvoyée à l'agent par le biais d'une récompense, positive ou négative selon l'effet sur l'environnement. Le nouvel état de ce dernier est également renvoyé à l'agent après son action. L'agent, à force d'interagir avec l'environnement, apprend à maximiser les récompenses qu'il obtient, et s'approche d'un schéma de décision optimal. Nous pouvons citer comme exemples d'applications l'optimisation de stratégie dans les jeux vidéos, la robotique, la reconnaissance d'objet ou d'image, etc.

### Définitions 6.1.1 -

Les notions centrales en Reinforcement Learning sont les suivantes :

- i. Un **agent** est un système capable de prendre des décisions en autonomie.
- ii. L'**environnement** constitue le monde dans lequel l'agent évolue et avec lequel il interagit.
- iii. L'agent perçoit à chaque instant  $t \in \mathbb{N}$ , l'**état**  $S_t$  de l'environnement, correspondant à la description de l'environnement.
- iv. A un instant  $t \in \mathbb{N}$ , l'agent interagit avec l'environnement en effectuant une **action**  $A_t$  qu'il décide.
- v. En fonction de la qualité de l'action effectuée sur l'environnement, une **récompense**  $R_t$  est rendue à l'agent.  
Nous notons par la suite  $\mathcal{A}$  l'ensemble des actions possibles,  $\mathcal{S}$  l'ensemble des états possibles et  $\mathcal{R}$  celui des récompenses possibles, de sorte que pour tout  $t \in \mathbb{N}$ ,  $A_t \in \mathcal{A}$ ,  $S_t \in \mathcal{S}$  et  $R_t \in \mathcal{R}$ .
- vi. La stratégie interne à l'agent qui lui permet de prendre des décisions est appelée **policy**, et est notée  $\pi_t : \mathcal{A} \rightarrow \mathcal{R}$  à l'instant  $t \in \mathbb{N}$ .

Les variables  $S_t, A_t, R_t$  sont des variables aléatoires qui se réalisent à chaque instant.

En résumé, l'agent est un système capable de prendre des décisions et mener des actions dans un environnement dans le but d'optimiser les récompenses obtenues. D'un point de vue probabiliste, la policy est une fonction qui donne la probabilité que l'agent a de faire une action  $a$  étant donné l'état  $s$  actuel de l'environnement. On note  $\pi_t(s, a) = \mathbb{P}(A_t = a | S_t = s)$ . Finalement, et d'après les définitions, l'objectif du R.L. est d'optimiser la stratégie opérée par l'agent pour maximiser les récompenses.

Cet agent prend donc ses décisions à des instants notés  $t \in \mathbb{N}$  du temps. Pour l'apprentissage par renforcement, le cadre épisodique est généralement adopté. Dans un cadre épisodique, nous considérons des séquences de temps  $t$  qui terminent presque sûrement. Le début d'un épisode est marqué par l'instant  $t = 0$ , avec donc  $S_0$  l'état initial de l'environnement. Prenons l'exemple des jeux vidéos, un épisode constitue une partie du jeu, dans lequel l'agent évolue jusqu'à la fin de la partie, que ce soit par la victoire ou la défaite. Une fois la partie finie, l'agent recommence une partie durant laquelle il a alors l'occasion de tester une nouvelle stratégie, d'améliorer la sienne, etc.



## *Q-fonction et équation de Bellman*

A chaque instant  $t \in \mathbb{N}$ , i.e. pour chaque épisode, l'agent récupère une récompense. Cette récompense doit être interprétée pour rendre compte de la qualité de ses décisions. L'instant auquel est pris l'action peut également être plus ou moins important pour l'action. Selon le cadre, l'environnement et le but, nous définissons le gain, censé rendre compte de l'influence d'une récompense locale sur l'ensemble de l'épisode. Pour la suite nous définissons le gain de la façon suivante :

Définition 6.1.2 -

Soit  $\gamma \in ]0; 1[$ , appelé *discount*, on définit le gain à partir du temps  $t \in \mathbb{N}$  comme :

$$G_t = \sum_{i=0}^{\infty} \gamma^i R_{t+i+1}$$

Dans le contexte des jeux vidéos par exemple, le gain représente donc le score atteint au cours de la partie, après un instant  $t$ . Ce gain permet de faire état de la réussite ou non de l'agent. Le but est à nouveau de le maximiser en adaptant les actions dans le but d'obtenir de meilleures récompenses à chaque instant. Il existe plusieurs manières d'améliorer cette stratégie, qui est donc comprise dans la policy  $\pi$  que suit l'agent, et nous définissons pour ce rapport la méthode suivante :

Définition 6.1.3 -

On appelle fonction action-valeur, *Q-value function* ou encore **Q-fonction**, noté couramment  $Q_{\pi,t}$  d'une policy  $\pi$  à l'instant  $t$  la fonction :

$$Q_{\pi,t} : (s, a) \in \mathcal{S} \times \mathcal{A} \mapsto \mathbb{E}_{\pi}[G_t | S_t = s, A_t = a]$$

On cherche alors à déterminer la *Q-value function* optimale  $Q^*(s, a) = \max_{\pi} Q_{\pi}(s, a)$ . Elle correspond au meilleur gain possible étant donné un état  $s$  et une action  $a$ . On définit de même  $\pi^*$  une stratégie telle que  $Q_{\pi^*}(s, a) = Q^*(s, a)$ . Par la suite, on notera  $Q$  plutôt que  $Q_{\pi}$  pour simplifier les notations.

Proposition 6.1.0 - Équation de Bellman

Soit un instant  $t \in \mathbb{N}$ , le gain vérifie :

$$G_t = R_{t+1} + \gamma G_{t+1} \quad (*)$$

De cette équation, nous pouvons exprimer alors récursivement la *Q-fonction*. Soit un état  $s$  et une action  $a$ , on déduit de (\*) :

$$Q(s, a) = \mathbb{E}_{s' \sim P(s'|s, a)} [r + \gamma \mathbb{E}_{a' \sim \pi} [Q(s', a')]] \quad (**)$$

où  $s'$  représente l'état au temps suivant, et  $r$  la récompense.

Introduisons l'opérateur de Bellman

$$\mathcal{T}^{\pi} : Q \mapsto ((s, a) \mapsto \mathbb{E}_{s' \sim P(s'|s, a)} [r + \gamma \mathbb{E}_{a' \sim \pi} [Q(s', a')]])$$

et similairement l'opérateur optimal de Bellman :

$$\mathcal{T}^* : Q \mapsto \left( (s, a) \mapsto \mathbb{E}_{s' \sim P(s'|s, a)} \left[ r + \gamma \max_{a' \in \mathcal{A}} Q(s', a') \right] \right)$$

Théorème 6.1.1 -

La *Q-fonction* optimale est l'unique point fixe de l'opérateur de Bellman optimal.

$Q^*$  peut ainsi être approchée itérativement en appliquant les formules théoriques précédentes. La formule de récurrence implémentable s'écrit alors :

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha \left( r_{t+1} + \gamma \max_{a \in \mathcal{A}} Q_t(s_{t+1}, a) - Q_t(s_t, a_t) \right)$$

avec  $\alpha$  le taux d'apprentissage. Il correspond au coefficient déterminant à quel point la valeur de l'action devrait changer à chaque itération. Lorsque  $\alpha = 0$ , on voit que rien ne change, tandis que si  $\alpha = 1$ , alors  $Q(S_t, A_t) \leftarrow R_t + \gamma \max_{a \in \mathcal{A}} Q(S_{t+1}, a)$ . C'est à partir de cette équation que sont construits les algorithmes de  $Q$ -Learning.

### 6.1.2. $Q$ -Learning

Pour l'intérêt de l'article [3], la performance de l'agent est évaluée à l'aide de la  $Q$ -value fonction. Le  $Q$ -Learning est un algorithme de Reinforcement learning visant à trouver la meilleure action à effectuer, étant donné l'état actuel  $s_t$  de l'environnement, en utilisant la  $Q$ -fonction, et notamment l'équation de Bellman (\*\*). Le fonctionnement général de l'agent est le suivant : étant donné un état  $s_t$  à l'instant  $t$ , l'agent prend la décision d'effectuer une action  $a_t$ , amenant l'environnement à l'état suivant  $s_{t+1}$  avec une récompense  $r_t$ . Chaque action apporte une récompense que l'agent va alors retenir afin d'apprendre. L'agent a donc besoin d'une mémoire d'une capacité défini  $M$  dans laquelle il va sauvegarder chaque étape  $(s_t, a_t, s_{t+1}, r_t)$ .

Ainsi, l'intérêt du R.L. est de déterminer la  $Q$ -fonction optimale en chaque couple  $(s, a)$ . On peut donc représenter la  $Q$ -fonction comme un tableau, couramment appelé  $Q$ -tableau, dans lequel l'agent associe à chaque  $s$  (en ligne par exemple) et chaque  $a$  (en colonne) la valeur de la  $Q$ -fonction approchée. Ce tableau doit alors être déterminé au mieux par l'agent pour garantir une bonne stratégie de résolution. Grâce à cela, l'agent est alors capable de décider une action plutôt qu'une autre.

Remarquons que, si on prend l'exemple d'un jeu Atari comme présenté dans l'article [3], les jeux sont représentés sur un écran de  $210 \times 60$  pixels, avec 128 couleurs possibles pour chaque pixel. Cela représenterait un  $Q$ -tableau de grandes dimensions à explorer.

Maintenant, nous sommes confrontés à un problème possible : l'agent pourrait trouver une solution locale et manquer d'explorer peut être des possibilités meilleures. Imaginons par exemple que notre jeu soit un labyrinthe, avec deux chemins possibles : un long et un court. L'agent veut bien évidemment quitter le labyrinthe mais aussi le quitter le plus rapidement possible. Selon la position initiale de l'agent, les premiers coups qu'il choisit, etc, il pourrait trouver le chemin le plus long comme solution en premier, et seulement celle-là. Pour éviter cela, on utilise une approche  $\varepsilon$ -greedy, avec  $\varepsilon \in ]0; 1]$ . Cela consiste à donner à l'agent une probabilité  $\varepsilon$  de faire une action aléatoire à chaque instant. En apportant de la stochasticité à l'algorithme, l'agent peut être amené à explorer des solutions ou, du moins, d'autres pistes qu'il aurait peut être manqué. Que celles-ci soient bonnes ou mauvaises, elles permettent à l'agent d'en apprendre d'autant plus sur l'environnement. Le paramètre  $\varepsilon$  est appelé probabilité d'exploration, et  $1 - \varepsilon$  est donc la probabilité d'exploitation.

De manière générale, lorsque l'on implémente un algorithme de  $Q$ -Learning, la probabilité  $\varepsilon$  diminue très faiblement au cours du temps, jusqu'à atteindre un seuil minimal  $\varepsilon_0 > 0$ . L'idée est que plus le temps avance, plus l'agent est capable d'exploiter correctement ce qu'il a acquis comme expérience.

Ce procédé est mis en place dans les algorithmes de  $Q$ -learning. Le principe général de ces algorithmes est présenté dans l'algorithme 1. L'idée donc d'un algorithme de  $Q$ -Learning est d'approcher au mieux la

$Q$ -fonction en utilisant l'action qui optimise actuellement le modèle. Une mise-à-jour dans l'algorithme suit [5] :

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha \left( R_t + \max_{a \in \mathcal{A}} \gamma Q(S_{t+1}, a) - Q(S_t, A_t) \right) \quad (\text{maj})$$

où  $\alpha$  représente le taux d'apprentissage et  $\gamma$  le facteur de discount évoqué précédemment.

---

**Algorithme 1** : Algorithme de  $Q$ -Learning général [5, 6]

---

**Data** :  $\alpha \in ]0; 1]$ ,  $\gamma \in ]0; 1]$ ,  $\varepsilon, \varepsilon_0 \in ]0; 1[$  ;

**pour** *épisode allant de 1 à N* **faire**

$s_t = \text{reset}()$  ;

**tant que** *l'épisode n'est pas terminé* **faire**

**si**  $u \sim \mathcal{U}([0; 1])$ ,  $u < \varepsilon$  : **alors**

            Faire une action  $a_t$  aléatoire dans  $\mathcal{A}$

**sinon**

$a_t = \arg \max_{a \in \mathcal{A}} Q(s_t, a)$  ;

$r_t, s_{t+1} = \text{current\_state}()$  ;

$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha (r_t + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t))$  ;

$s_t = s_{t+1}$  ;

$a_t = a_{t+1}$

**fin**

**fin**

**fin**

---

Dans l'article étudié [3], les auteurs proposent des extensions de cette méthode générale et construisent un algorithme d'apprentissage d'autant plus efficace. Nous étudierons en détail deux d'entre eux : le Deep  $Q$ -Learning (DQN), développé en 6.2.3, et le Double Deep  $Q$ -Learning (DDQN), développé en 6.2.4.

---

## 6.2. Deep Q-Learning

---

Dans cette partie, nous allons explorer différents exemples d'algorithmes de  $Q$ -Learning qui ont été proposés dans l'article [3], c'est à dire des algorithmes de Deep  $Q$ -Learning. Nous nous concentrerons particulièrement sur l'algorithme DQN classique et le DDQN. Ces algorithmes ont été testés sur un jeu implémenté via *python* correspondant à un jeu *Snake* simplifié. L'implémentation de cet environnement est expliquée et présentée dans l'annexe B.

Dans les parties qui suivent, les algorithmes mentionnés sont accessibles sur [https://github.com/aurianegbt/Reinforcment\\_Learning\\_DQN](https://github.com/aurianegbt/Reinforcment_Learning_DQN).

### 6.2.1. Le Deep Q-Learning et les réseaux de neurones

De manière générale, les algorithmes  $Q$ -Learning sont donc basés sur une représentation en tableau, comme évoqué précédemment. Celle-ci permet d'explorer la  $Q$ -fonction en associant à chaque couple état-action  $(s, a)$  une valeur approchée de celle-ci. Malheureusement, dans certains cas, ce tableau peut vite être de grande dimension. Si on prend un exemple classique d'algorithme de Reinforcement Learning, comme la reconnaissance d'image, le nombre de pixels présent sur une image atteint vite plusieurs milliers. Cela nécessiterait d'explorer un très grand nombre d'état, ce qui n'est pas forcément réalisable dans un temps raisonnable. C'est pour résoudre ce problème que l'on a recourt au Deep Learning, en remplaçant le tableau par une fonction approximant notre  $Q$ -fonction. Nous réglons alors le problème de dimension en utilisant ce qu'on appelle un réseau de neurones. Le principe est le suivant : on a un réseau de neurones, capable d'approcher notre fonction, et une mémoire de capacité limitée  $M$ . On va alors utiliser cette mémoire, et donc les expériences acquises, pour mettre à jour notre approximation de  $Q$  au fur et à mesure des épisodes.

Comme son nom l'indique, un réseau de neurones est l'association de plusieurs neurones. Maintenant, un neurone est constitué d'une entrée, d'une sortie, respectivement analogues aux dendrites et à l'axone d'un neurone biologique. Entre les deux, le neurone effectue un calcul, une transformation, etc. On peut schématiser cela comme représenté figure 6.3.

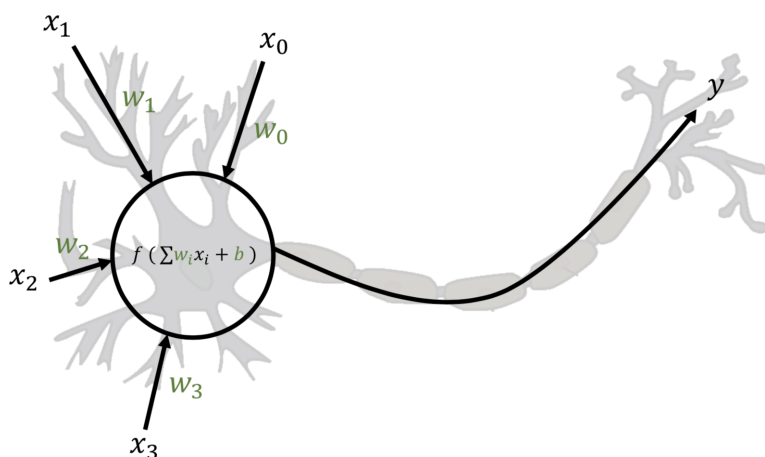


FIGURE 6.3 – Représentation schématique d'un neurone.

source : <https://culturesciencesphysique.ens-lyon.fr/>

On observe plusieurs entrées  $x_i$  qui sont pondérés par  $w_i$ , puis transformés selon une formule  $f : (x_i) \mapsto f\left(\sum_i x_i + b\right)$  définie pour le neurone, avec  $f$  une fonction croissante appelée une fonction d'activation. Les poids  $\theta = (w_i)$  correspondent au poids du neurone. Ces derniers sont modifiés au fur et à mesure de l'apprentissage pour corriger notre approximation.

Ces neurones sont donc liés, en ce qu'on appelle des couches, formant une architecture algorithmique complète. La première couche est généralement constituée d'un certain nombre de neurones qui prennent en entrée les données brutes. Une deuxième couche de neurones va alors prendre en entrée toutes les sorties de la première couche et les transformer à nouveau et ainsi de suite jusqu'à la dernière couche. La dernière couche renvoie enfin une sortie qui va déterminer la réponse de l'algorithme par rapport à

nos entrées initiales.

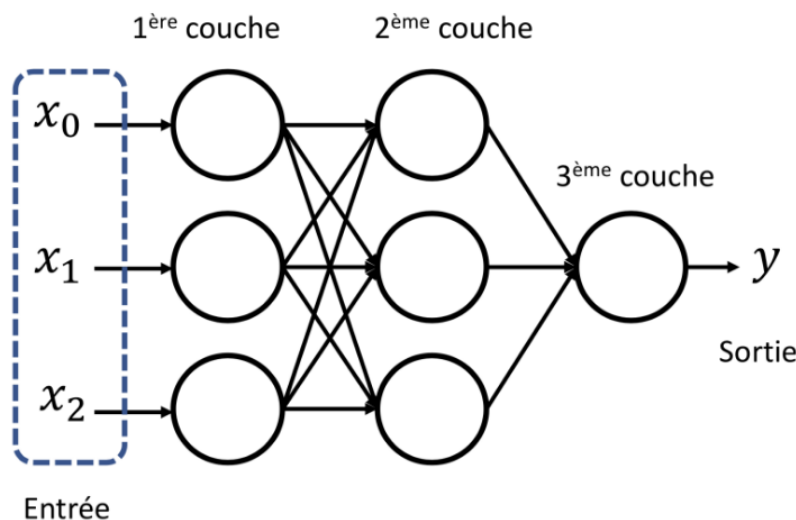


FIGURE 6.4 – Schéma d'un réseau de neurone.

source : <https://culturesciencesphysique.ens-lyon.fr/>

Lorsque l'on entraîne un agent, à chaque instant, on récupère la récompense et l'issue de l'action. On récupère donc un aperçu de la qualité de notre action. Nous avons en notre possession la connaissance de la sortie et de l'entrée de notre réseau. Nous sommes alors capable de corriger les poids ( $w_i$ ) du réseau pour faire correspondre au mieux nos entrées à notre sortie [7] .

### 6.2.2. Le *Q-Network*

Comme expliqué en partie 6.1.2, l'algorithme Deep *Q*-Network (DQN) est un algorithme de *Q*-Learning, utilisant un réseau de neurones. Nous avons évoqué l'utilisation de *Q*-table qui pouvait, dans certains environnements, être remplacé par un réseau de neurones. Ceci est le principe de l'algorithme DQN . Plutôt que de représenter notre approximation de la *Q*-fonction par un tableau, on utilise un réseau de neurones pour gagner en efficacité et en rapidité. Nous allons détailler les composants de notre réseau, à savoir ce qu'on a en entrée, ce qu'on a en sortie et l'architecture entre [5, 8].

#### Les entrées

En entrée, il faut fournir l'état dans lequel on se situe, ainsi que les positions actuelles des éléments sur le jeu. Si nous citons l'exemple des jeux Atari utilisés par les auteurs de l'article que l'on étudie, le jeu est représenté sur une image de plusieurs milliers de pixels. Sur cette image, on peut situer les éléments du jeu.

Dans notre exemple de *Snake*, l'image permet de situer le serpent, et la pomme. Mais cette image à l'instant  $t$  seule ne suffit pas pour décrire l'état du jeu actuel, on ne saurait déterminer dans quelle direction va le serpent, quel bout représente la tête et lequel la queue. Pour remédier à cela, on retient comme état du jeu à l'instant  $t$  les  $n_{last\_actions} > 1$  images du jeu. Cela suffit pour déduire où se dirige le serpent.

Dans cette configuration, si la grille de jeu mesure  $n_{row}$  unités de hauteur et  $n_{col}$  unités de largeur, il faut sauvegarder l'état  $s_t$  comme un tableau de dimension  $(n_{row}, n_{col}, n_{last\ actions})$ .

### La sortie

Ce que l'on veut récupérer à la fin des calculs, c'est une action à effectuer. Nous souhaitons savoir quelle action faire étant donné l'état dans lequel on est (i.e. les entrées fournies). Cette action est celle maximisant la  $Q$ -fonction dans l'état actuel, i.e. :

$$a_t = \arg \max_{a \in \mathcal{A}} Q(s_t, a)$$

Pour déterminer cette action, il faut donc calculer les différentes valeurs de la  $Q$ -fonction en chaque  $a \in \mathcal{A}$ . C'est donc ceci que l'on va demander aux neurones : la sortie du réseau de neurones doit être la valeur de la  $Q$ -fonction en l'ensemble des actions  $a \in \mathcal{A}$ .

### Le Modèle

D'après ce que l'on vient d'expliquer, le réseau de neurones utilisé serait chargé de calculer la  $Q$ -fonction en l'état  $s_t$  actuel et en chaque  $a \in \mathcal{A}$ . Cela n'est pas faisable directement et on a recourt à une approximation de  $Q$  que l'on note  $Q_\theta$ , où  $\theta$  représente les poids du réseau de neurones. A mesure que l'algorithme s'entraîne et progresse, il met à jour les poids  $\theta$  du réseau de neurones pour fournir une approximation plus précise de  $Q$ . Finalement, notre réseau en lui même applique à notre entrée  $s_t$  la fonction  $Q_\theta(s_t, \cdot)$  en chaque  $a \in \mathcal{A}$ . On utilisera comme réseau de neurones un réseau de neurones convolutif, qui est mieux adapté pour traiter en entrée des images [4, 5, 7, 9].

### 6.2.3. Mise en place de l'algorithme DQN

L'algorithme Deep  $Q$ -Network est composé de différentes phases. On a d'abord la phase d'initialisation, c'est simplement l'initialisation de l'environnement, de la mémoire, du réseau de neurones, etc. Ensuite, nous allons débiter l'entraînement de l'agent. Pour cela, on joue un grand nombre  $N$  d'épisodes, ou bien, dans le cas où il s'agit d'un jeu vidéo, un grand nombre  $N$  de parties. Au cours d'une partie, l'agent va effectuer un certain nombre d'actions, qu'il va sauvegarder en mémoire pour permettre de s'améliorer ensuite. Pour apprendre de son expérience, l'algorithme DQN utilise des mini-batches de mémoire, avec lesquels on met à jour les poids  $\theta$  du réseau de neurones [6, 8, 10].

### Entraînement de l'agent

Nous allons d'abord nous intéresser à l'apprentissage de l'agent et, en particulier, de quelle manière, dans les faits, il apprend d'une expérience. On a déjà mentionné précédemment le recours à une mémoire. Cette mémoire à une capacité bornée  $M$  assez grande (*une longueur de l'ordre de 50000 par exemple et dans notre cas*). Cette mémoire est une liste, contenant les  $M$  dernières expériences faites ; lorsque le temps progresse, il faut supprimer les plus anciennes expériences sauvegardées pour enregistrer les nouvelles. Chaque élément de cette mémoire correspond à un tuple  $(s_t, a_t, r_t, s_{t+1})$ . Cette mémoire va être utilisée pour mettre à jour les poids du réseau de neurones en prélevant des mini-batches d'expérience.

| Paramètres                         | Valeurs                   |
|------------------------------------|---------------------------|
| la capacité de la mémoire          | $M = 50000$               |
| la taille des batchs               | $b = 32$                  |
| le taux d'apprentissage            | $\alpha = 0.0001$         |
| le facteur de discount             | $\gamma = 0.9$            |
| le seuil du taux d'exploration     | $\varepsilon_0 = 0.05$    |
| Taux de décroissance d'exploration | $\varepsilon_- = 0.0002.$ |

TABLE 6.1 – Hyperparamètres utilisés

C'est à partir de ces mini-batchs que l'on entraîne notre agent en utilisant directement les méthodes proposées par le package `tensorflow`. Le principe est le suivant : on prélève aléatoire  $m$  expériences passés avec leur issue connue. Ces expériences sont des tuples  $(s^i, a^i, r^i, s_+^i)$  auxquels on associe

$$y_i = r^i + \gamma \max_{a_+^i} Q_\theta(s_+^i, a_+^i) \quad (6.1)$$

où les quantités  $x_+$  représentent la quantité  $x$  au temps suivant celui de l'expérience. Ceci correspond à la  $Q$ -value cible de chaque expérience, et on souhaite que notre  $Q_\theta$  courant l'approche au mieux. On met donc finalement à jour les poids de notre réseau de neurones en optimisant par descente de gradient avec une perte des moindres carrés, par exemple, pour trouver  $\theta$  minimisant la fonction de coût

$$L : \theta \mapsto \mathbb{E}[(y_i - Q_\theta(s_i, a_i))^2]$$

### Implémentation de l'algorithme

Pour implémenter l'algorithme, nous définissons des hyperparamètres résumés dans le tableau 6.1. Ces hyperparamètres ont été choisis à l'aide des références citées tout au long de ce rapport. L'implémentation de l'algorithme DQN suit celui de l'algorithme 2, qui reprend les différentes parties expliquées jusque là.

---

**Algorithme 2** : Algorithme Deep  $Q$ -Learning [5, 6]
 

---

**Data** : On initialise les poids  $\theta$  des modèles, la mémoire  $\mathcal{M}$  de capacité  $M$ , l'environnement  $env$  et un agent  $agent$  ;

**tant que** *On continue d'entraîner l'algorithme faire*

  On réinitialise  $env$ ,  $s_t$ ,  $s_{t+1}$  ;  $episode = 0$  ;  $\varepsilon = 0$  ;  $\varepsilon_-$  ;

$game\_over = \text{False}$  ;

**tant que**  $game\_over = \text{False}$  **faire**

**si**  $u \sim \mathcal{U}([0; 1])$ ,  $u < \varepsilon$  **alors**

      Choisir une action  $a_t$  aléatoire dans  $\mathcal{A}$  ;

**sinon**

      Choisir  $a_t$  maximisant  $Q_\theta(s_t, \cdot)$  ;

**fin**

    On garde  $s_{aux} = s_{t+1}$  ;

$agent$  effectue l'action  $a_t$  et récupère  $r_t$  et  $s_{t+1}$ , et l'état de  $game\_over$  ;

    On ajoute  $(s_t, a_t, r_t, s_{t+1}, game\_over)$  à  $\mathcal{M}$  ;

    On prélève un mini-batch  $\mathcal{B} = \{(s^i, a^i, r^i, s_+^i)\}_{i \leq b}$  d'expériences de  $\mathcal{M}$  ;

    On entraîne le modèle sur  $\mathcal{B}$  et met à jour les poids : **pour**  $i \leq b$  **faire**

**si**  $game\_over$  **alors**

$y^i = r^i$

**sinon**

$y^i = r^i + \gamma \max_{a_+} Q_\theta(s_+^i, a_+)$

**fin**

**fin**

    On met à jour les poids  $\theta$  par descente de gradient avec la perte  $L$ .

**fin**

  On met à jour  $episode$ ,  $\varepsilon$ ,  $s_t$ ,  $s_{t+1}$

**fin**

---

Nous avons implémenté et testé l'algorithme 2 sur 4 fichiers distincts, accessibles sur [https://github.com/aurianegbt/Reinforcment\\_Learning\\_DQN](https://github.com/aurianegbt/Reinforcment_Learning_DQN). :

- i. `environment.py`, implémentant le jeu de *Snake*, et dont le code est commenté partie B ;
- ii. `cnn.py`, correspondant à notre modèle de couches de neurones, vu dans la partie 6.2.1 ;
- iii. `DQN.py`, qui correspond à notre agent, doté d'une mémoire ;
- iv. `train.py`, il s'agit de l'algorithme 2 implémenté.

Après avoir lancé l'entraînement, celui-ci a été stoppé le plus tardivement possible. Le score, ici, correspond au nombre de pomme obtenue au cours d'une partie ; un score de 5 pour un épisode équivaut donc à une partie avec un score final de 5 pommes récupérées par le serpent avant que celui-ci meurt.



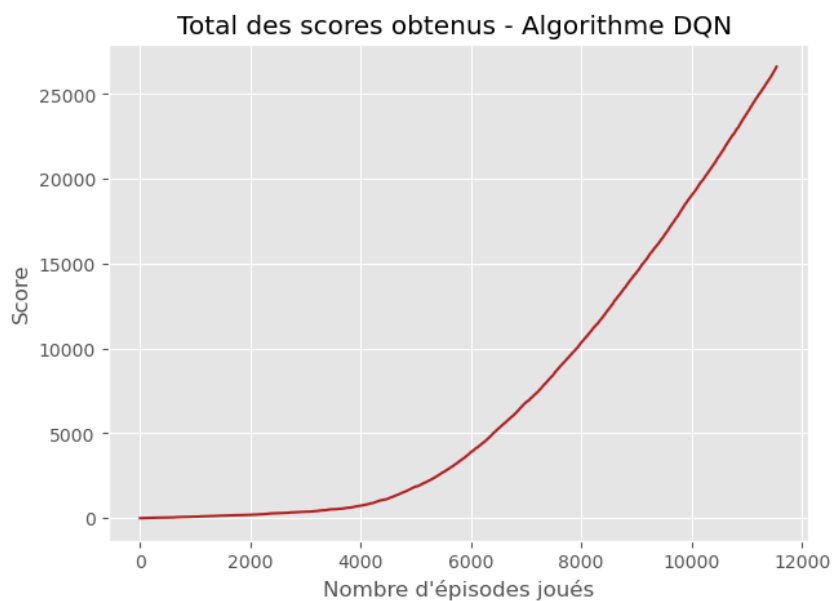


FIGURE 6.5 – Score total obtenu au fur et à mesure du temps par l’agent.

Nous appelons par la suite score le nombre de pommes gagnées en une partie. Comme on peut voir figure 6.5, le score total obtenu par l’agent ne cesse de croître avec une pente qui augmente d’abord puis semble tendre vers une droite linéaire. On peut observer trois phases à l’apprentissage de notre agent. Pendant la première phase, ou phase initiale, le score de l’agent reste bas, la plupart des parties se terminent avec un score nul (87%), et le score moyen obtenu n’est que de 0,15 tout le long de la phase. Dans la seconde phase, l’agent progresse rapidement, passant d’un score moyen de 0,26 à 4,36 en au moins 6400 épisodes. Au cours de cette deuxième phase, le pourcentage de parties aboutissant à un score nulle n’est plus que de 13%, et celui de parties avec un score inférieur à 53%. Enfin, on constate une dernière phase où l’agent semble avoir appris tout ce qu’il pouvait, il atteint un score maximale moyen de 4,94. Ce seuil maximal peut dépendre de la capacité de la mémoire fournie à l’agent notamment. Au cours de cette dernière phase, l’agent obtient sur 77% de ses parties un score supérieur à 3 et sur 53%

un score supérieur à 5.

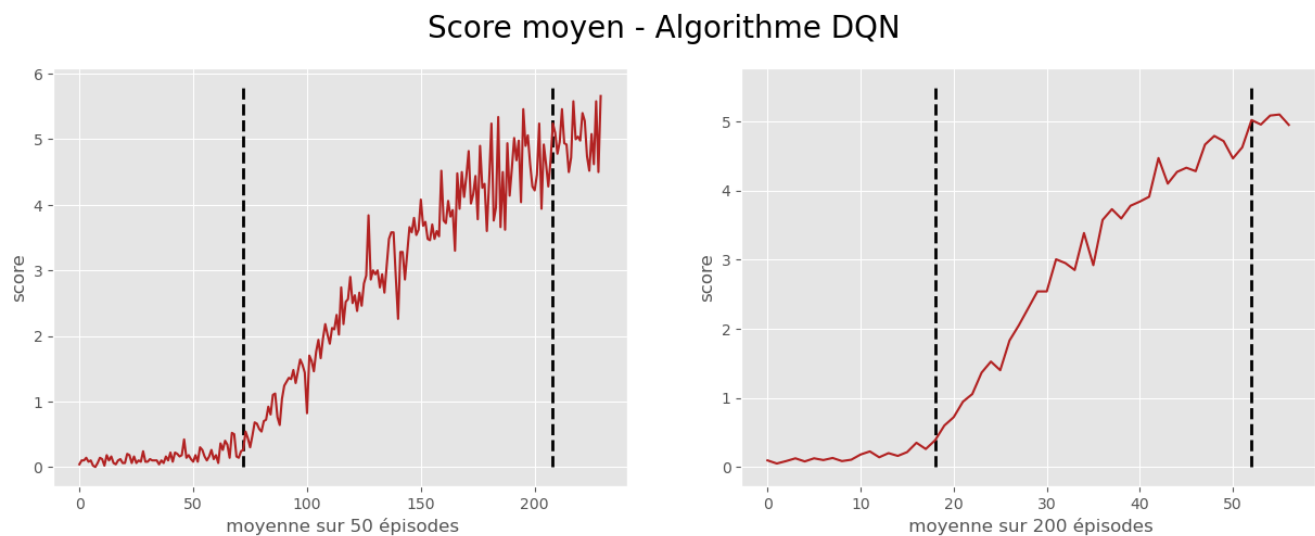


FIGURE 6.6 – Score moyen obtenu par l’agent au cours des  $N$  derniers épisodes, avec  $N = 50$  et 200.

#### 6.2.4. Double $Q$ -Learning

Nous allons maintenant explorer une des extensions de l’algorithme DQN proposé dans l’article [3], à savoir l’algorithme DDQN [5, 6, 11].

##### Principe

L’algorithme DDQN, pour Double Deep  $Q$ -Learning, répartit la  $Q$ -fonction sur deux modèles :

- un premier, appelé *target-network* usuellement, qui s’occupe spécifiquement de l’estimation de la  $Q$ -fonction, noté  $Q^1$  ;
- un second spécifiquement dédié au choix de l’action, i.e. l’action maximisant la  $Q$ -fonction, noté  $Q^2$ .

Ces changements transforment la formule approchant la  $Q$ -fonction de l’algorithme DQN 6.1 en :

$$y_i = r^i + \gamma Q^1(s_+^i, \arg \max_{a_+^i} Q^2(s_+^i, a_+^i)) \quad (6.2)$$

Du fait de cette construction, nous avons deux poids  $\theta^1$  et  $\theta^2$  à mettre à jour, mais pour éviter tout problème de divergence des modèles, les poids sont régulièrement harmonisés.

##### Implémentation

Lors de l’implémentation de l’algorithme DDQN, on reprend exactement les hyperparamètres du tableau 6.1, pour pouvoir comparer les performances des deux algorithmes ensuite.

Nous avons implémenté, de la même manière que précédemment, l’algorithme sur 4 fichiers distincts, accessibles sur [https://github.com/aurianegbt/Reinforcement\\_Learning\\_DQN](https://github.com/aurianegbt/Reinforcement_Learning_DQN). :

**Algorithme 3** : Algorithme Double Deep  $Q$ -Learning [5, 6]

---

**Data** : On initialise les poids  $\theta^1 = \theta^2$  des modèles, la mémoire  $\mathcal{M}$  de capacité  $M$ , l'environnement  $env$  et un agent  $agent$  ;

**tant que** *On continue d'entraîner l'algorithme* **faire**

$env = reset()$  ;  $s_t, s_{t+1} = reset()$  ;  $episode = 0$  ;  $\varepsilon = 0$  ;  $\varepsilon_-$  ;

$game\_over = \mathbf{False}$  ;

**tant que**  $game\_over = \mathbf{False}$  **faire**

Avec probabilité  $\varepsilon$ , choisir  $a_t$  aléatoirement dans  $\mathcal{A}$ , sinon  $a_t \in \arg \max_{\mathcal{A}} Q_{\theta^2}^2(s_t, \cdot)$  ;

$agent$  effectue l'action  $a_t$  et récupère  $r_t$  et  $s_{t+1}$ , et l'état de  $game\_over$  ;

On ajoute  $(s_t, a_t, r_t, s_{t+1}, game\_over)$  à  $\mathcal{M}$  ;

On prélève un mini-batch  $\mathcal{B} = \{(s^i, a^i, r^i, s_+^i)\}_{i \leq b}$  d'expériences de  $\mathcal{M}$  ;

On entraîne le modèle sur  $\mathcal{B}$  : **pour**  $i \leq b$  **faire**

**si**  $game\_over$  **alors**

$y^i = r^i$

**sinon**

$y^i = r^i + \gamma Q_{\theta^1}^1(s_+^i, \arg \max_{a \in \mathcal{A}} Q_{\theta^2}^2(s_t, a))$

**fin**

**fin**

On effectue une descente de gradient avec la perte  $L$  sur  $\theta_1$  ;

On met à jour les poids  $\theta_1$  et, tout les  $K$  épisodes,  $\theta_2$  aussi ;

**fin**

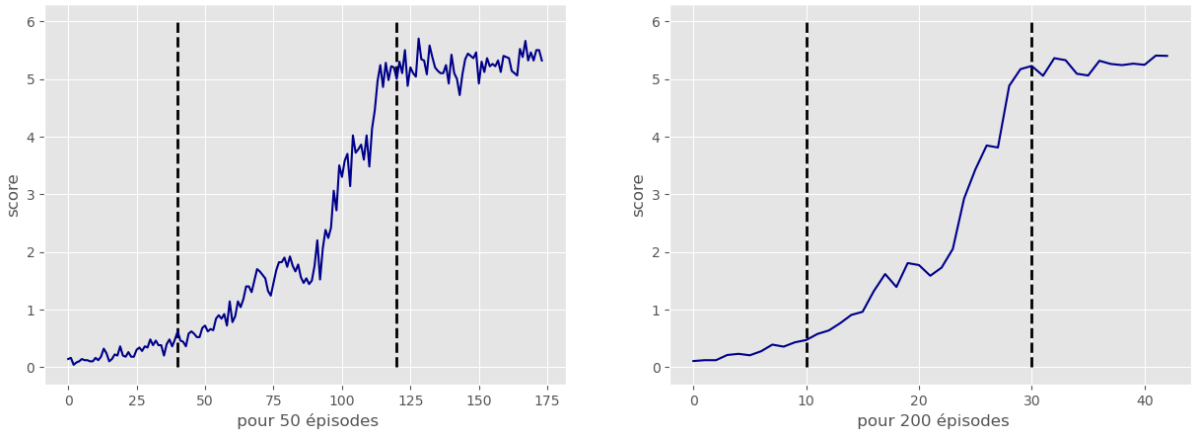
On met à jour  $episode$ ,  $\varepsilon$ ,  $s_t$ ,  $s_{t+1}$

**fin**

---

- i. `environment.py` de nouveau ;
- ii. `cnn.py`, correspondant à notre modèle de couches de neurones, vu dans la partie 6.2.1 ;
- iii. `DDQN.py`, qui correspond à notre agent ;
- iv. `train.py`, l'algorithme 3 donnée.

Score moyen - Algorithme DDQN

FIGURE 6.7 – Score moyen obtenu par l'agent au cours des  $N$  derniers épisodes, avec  $N = 50$  et 200.

On observe la même progression que lors de l'algorithme DQN, avec un apprentissage en trois phases, comme représentés sur les figures 6.7 et 6.8. L'agent est d'abord dans une phase initiale avec un score faible, 0.24 en moyenne, pendant 2000 épisodes. Au cours de cette phase, l'agent termine 78% de ses

parties avec un score nul. L'agent progresse ensuite rapidement et passe d'un score de 0.6 pommes en moyenne à 5.2 en 4000 épisodes environ. Cette fois ci, le pourcentage de parties terminées avec un score nul n'est plus que de 20%, et le pourcentage de parties avec un scores inférieur à 3 est de 67%. Le seuil maximal de l'algorithme DDQN implémenté semble alors être atteint au bout de 6000 épisodes, et vaut 5.26 environ. Dans cette dernière phase, l'agent obtient plus de 5 pommes dans 75% des parties, et plus de 3 dans 93% des parties.

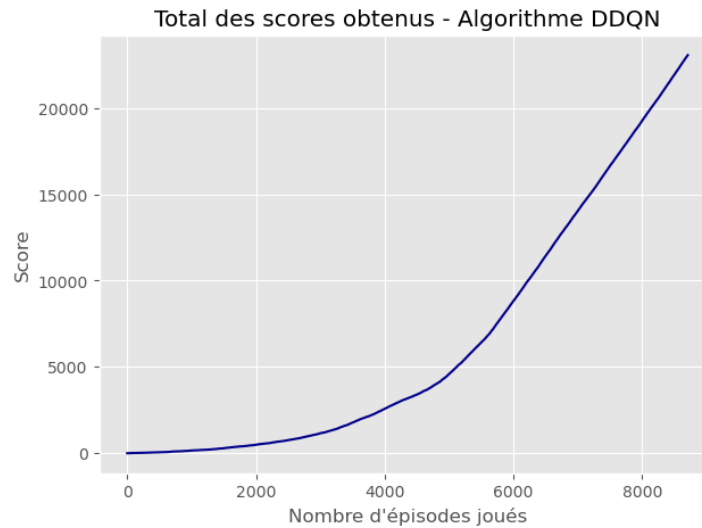


FIGURE 6.8 – Score total obtenu au fur et à mesure du temps par l'agent.

#### DQN vs DDQN

Les deux algorithmes présentés ont été lancés sur deux ordinateurs différents, donc nous ne pouvons que partiellement commenter le temps d'exécution. Les deux algorithmes ont commencé au même rythme (30sec par épisode). L'algorithme DDQN n'a pas significativement ralenti au fur et à mesure du temps, tandis que celui DQN a largement ralenti, passant d'environ 30sec par épisode à presque 5minutes sur les 3000 derniers. Bien que sur des supports différents, le ralentissement observé dans le DQN semble très marqué.

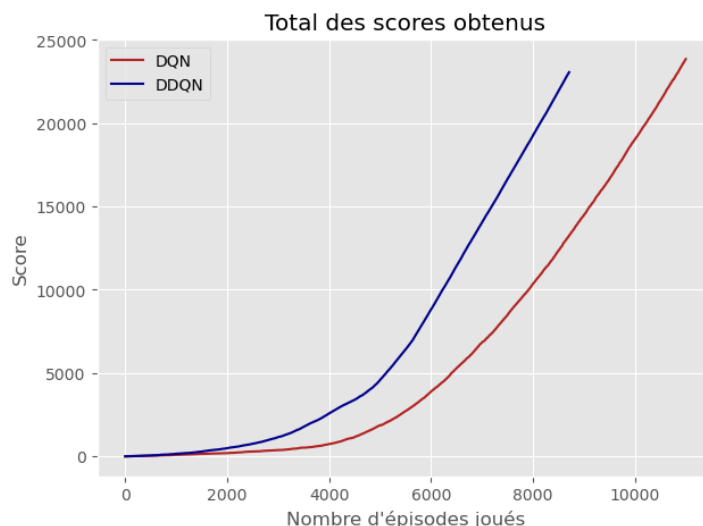
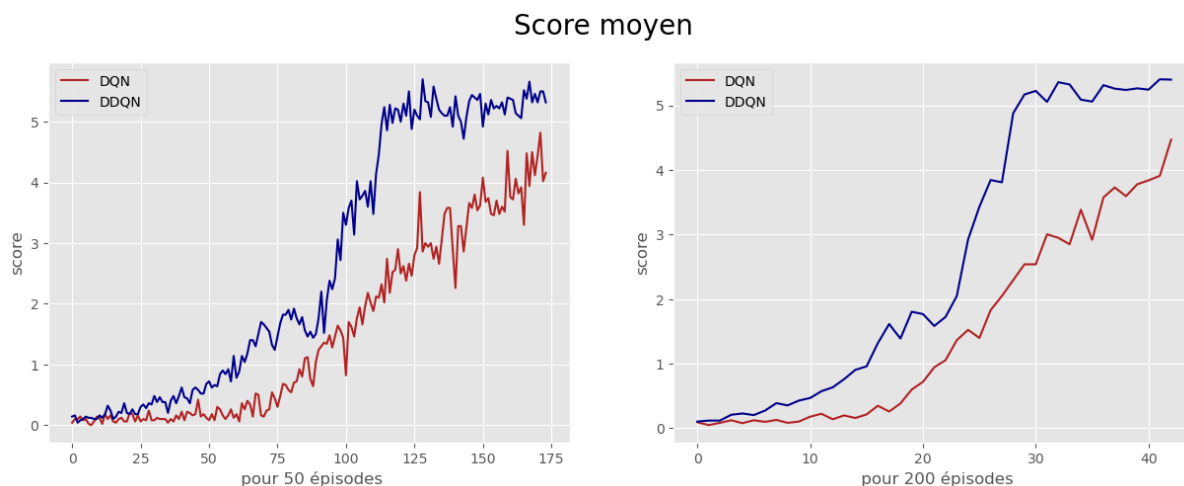


FIGURE 6.9 – Score total obtenu au fur et à mesure du temps par l'agent DQN et DDQN

En terme de performance, l'algorithme DDQN parvient à doubler le score moyen de celui DQN, comme visible sur la figure 6.9. Le score final de l'algorithme DDQN au 8000<sup>e</sup> épisode est de 19262 pommes contre 10358 pour le DQN, soit 46% de moins. Dans la même idée, nous observons sur la figure 6.10 que la première phase est beaucoup moins marquée chez l'algorithme DDQN que chez le premier, avec une progression existante, bien que moins marquée que dans ce qu'on a appelé la deuxième phase. Enfin, pour atteindre ce que nous avons appelé la dernière phase, il faut 6000 épisodes pour l'algorithme DDQN contre 10000 pour celui DQN.

FIGURE 6.10 – Score moyen obtenu par l'agent DQN et DDQN au cours des  $N$  derniers épisodes, avec  $N = 50$  et 200

Enfin, nous pouvons également remarquer les problèmes de stabilités mentionnés ultérieurement. Sur le graphique à gauche de la figure 6.10 de gauche, le score oscille plus chez le DQN que le DDQN, et admet des pics bien discernables et marqués, avec des scores plus bas de  $\pm 1$  pomme par rapport à la moyenne locale (sur 200 épisodes comme représenté à droite).

Il a été montré [5, 6, 11] que cette extension de l'algorithme DQN répondait au problème de sures-

timisation de la  $Q$ -fonction par l'algorithme classique, et fournissait des résultats plus rapidement que le premier.

---

## 6.3. Algorithme Rainbow

---

Nous allons maintenant passer en revue les autres extensions utilisées pour construire l'algorithme "Rainbow" [3].

### 6.3.1. Prioritized Replay

L'algorithme Prioritized Experience Replay (PER) DQN est une des extensions mentionnées dans le papier [3] étudié. La différence avec l'algorithme DQN est les mini-batches utilisés pour entraîner notre modèle. Plutôt que de choisir aléatoirement et uniformément parmi les expériences en mémoire, on les choisit aléatoirement selon un ordre de priorité. Certaines expériences stockées en mémoire par l'agent vont être décrétées plus importantes, donc avoir une priorité  $p_i$  plus importante. A chaque expérience dans la mémoire est alors associée une priorité  $p_i \in ]0; 1[$ ,  $i \leq M$ . Une expérience de priorité  $p_i$  a donc une probabilité  $\frac{p_i}{\sum_j p_j}$  d'être sélectionnée lors de la sélection d'un mini-batch. Les expériences vont ensuite être mis-à-jour au fur et à mesure de l'entraînement comme suit :

- Une nouvelle expérience va avoir sa priorité initialisée au max des priorités déjà définies,
- Lorsque l'on entraîne le modèle, on calcule pour chaque élément du mini-batch son erreur par rapport à la valeur cible, cette valeur va remplacer la priorité de l'expérience.

Le choix de cette mis-à-jour vient du fait que l'expérience semble d'autant plus inattendue que cette erreur est élevée. Il semble donc que l'agent puisse apprendre d'autant plus de cette expérience que d'une plus ordinaire parmi les éléments de la mémoire. Les changements mentionnés sont présentés dans l'algorithme modifié 4.

Cette modification peut aider à accélérer la convergence de l'algorithme vers la  $Q$ -fonction et améliorer plus rapidement les performances de l'algorithme [12].

### 6.3.2. Dueling Networks

Le *Dueling Network* est, cette fois, une extension du DDQN plutôt. Il reprend le principe de l'algorithme DDQN en ajoutant une architecture spécifique au réseau de neurones  $Q^2$  utilisée pour prédire la valeur d'une action à partir d'un état donné. Il consiste à séparer les calculs en une partie "valeur" et

**Algorithme 4 :** Algorithme Double Deep  $Q$ -Learning avec priorité [12]

**Data :** On initialise les poids  $\theta^1 = \theta^2$  des modèles, la mémoire  $\mathcal{M}$  de capacité  $M$ , l'environnement  $env$  et un agent  $agent$ .

**tant que** *On continue d'entraîner l'algorithme faire*

$env = reset(); s_t, s_{t+1} = reset(); episode = 0; \varepsilon = 0; \varepsilon_-;$

;

$game\_over = \mathbf{False};$

;

**tant que**  $game\_over = \mathbf{False}$  **faire**

Avec probabilité  $\varepsilon$ , choisir  $a_t$  aléatoirement dans  $\mathcal{A}$ , sinon  $a_t \in \arg \max_{\mathcal{A}} Q^2(s_t, \cdot);$

$agent$  effectue l'action  $a_t$  et récupère  $r_t$  et  $s_{t+1}$ , et l'état de  $game\_over$ ;

On ajoute  $(s_t, a_t, r_t, s_{t+1}, game\_over)$  à  $\mathcal{M}$  avec la priorité  $p_t = \max_{i < t} p_i$ ;

On prélève un mini-batch  $\mathcal{B} = \{(s^i, a^i, r^i, s_+^i)_{i \leq b}\}$  d'expériences de  $\mathcal{M}$  selon les priorités  $(p_i)_{i \leq M}$ ;

On entraîne le modèle sur  $\mathcal{B}$  : **for**  $i \leq b$  **do**

**si**  $game\_over$  **alors**

$y^i = r^i$

**sinon**

$y^i = r^i + \gamma Q_{\theta^1}^1(s_+^i, \arg \max_{a \in \mathcal{A}} Q_{\theta^2}^2(s_t, a))$

**fin**

On calcule l'erreur  $\delta_i = |y_i - Q_{\theta^1}^1(s_+^i, a_+^i)|$  et remplace  $p_i$ ;

**end**

On effectue une descente de gradient avec la perte  $L$  sur  $\theta_1$ ;

On met à jour les poids  $\theta_1$  et, tout les  $K$  épisodes,  $\theta_2$  aussi;

**fin**

On met à jour  $episode, \varepsilon, s_t, s_{t+1}$

**fin**

une partie “avantage”, sur deux couches distinctes du réseau de neurones mais liées. Le découpage suit :

$$Q^2(s, a) = \underbrace{\mathbb{E}_{a \sim \pi(s)}[Q^2(s, a)]}_{=V(s)} + \underbrace{Q^2(s, a) - \mathbb{E}_{a \sim \pi(s)}[Q^2(s, a)]}_{=A(s, a)}$$

où  $V$  représente la première fonction “valeur”, et  $A$  est la fonction “avantage”.

---

**Algorithme 5** : Algorithme Dueling Deep  $Q$ -Learning [13]
 

---

**Data** : On initialise les poids  $\theta^1$ ,  $\alpha$  et  $\beta$  des modèles, la mémoire  $\mathcal{M}$  de capacité  $M$ , l’environnement *env* et un agent *agent*.

**tant que** *On continue d’entraîner l’algorithme faire*

*env* = *reset*();  $s_t, s_{t+1}$  = *reset*(); *episode* = 0;  $\varepsilon$  = 0;  $\varepsilon_-$ ;

;

*game\_over* = **False**;

**tant que** *game\_over* = **False** **faire**

Avec probabilité  $\varepsilon$ , choisir  $a_t$  aléatoirement dans  $\mathcal{A}$ ;

sinon  $a_t \in \arg \max_{\mathcal{A}} \left[ V_{\theta, \alpha}(s) + \left( A_{\theta, \beta}(s, a) - \frac{1}{\text{card } \mathcal{A}} \sum_{a' \in \mathcal{A}} A_{\theta, \beta}(s, a') \right) \right]$ ;

*agent* effectue l’action  $a_t$  et récupère  $r_t$  et  $s_{t+1}$ , et l’état de *game\_over*;

On ajoute  $(s_t, a_t, r_t, s_{t+1}, \text{game\_over})$  à  $\mathcal{M}$ ;

On prélève un mini-batch  $\mathcal{B} = \{(s^i, a^i, r^i, s_+^i)_{i \leq b}\}$  d’expériences de  $\mathcal{M}$ ;

On entraîne le modèle sur  $\mathcal{B}$  : **for**  $i \leq b$  **do**

**si** *game\_over* **alors**

$y^i = r^i$

**sinon**

$y^i = r^i + \gamma Q_{\theta^1}^1(s_+^i, \arg \max_{a \in \mathcal{A}} V_{\theta, \alpha}(s) + \left( A_{\theta, \beta}(s, a) - \frac{1}{\text{card } \mathcal{A}} \sum_{a' \in \mathcal{A}} A_{\theta, \beta}(s, a') \right))$

**fin**

**end**

On effectue une descente de gradient avec la perte  $L$  sur  $\theta_1$ ;

On met à jour les poids  $\theta_1$  et, tout les  $K$  épisodes,  $\alpha, \beta$  aussi;

**fin**

On met à jour *episode*,  $\varepsilon$ ,  $s_t$ ,  $s_{t+1}$

**fin**

---

Néanmoins, ce découpage présente un problème : l’unicité pour une fonction  $Q^2$  donné du couple  $(V, A)$ . Une manière de répondre à ce problème est d’imposer une certaine valeur à l’une des deux fonctions. Le choix fait par les auteurs de l’article [13] sur lequel est basé celui que l’on étudie [3] est donc de considérer :

$$Q^2(s, a) = V(s) + \left( A(s, a) - \frac{1}{\text{card } \mathcal{A}} \sum_{a' \in \mathcal{A}} A(s, a') \right)$$

Étant implémentée sur deux couches différentes, la fonction  $V$  retournée par le réseau de neurones utilisé dépend des poids  $\alpha$  de ce dernier, distincts des poids  $\beta$  de celui de  $A$ . Nous pouvons finalement écrire :

$$Q_{\theta}^2(s, a) = V_{\theta, \alpha}(s) + \left( A_{\theta, \beta}(s, a) - \frac{1}{\text{card } \mathcal{A}} \sum_{a' \in \mathcal{A}} A_{\theta, \beta}(s, a') \right)$$



Les changements sont rapportés sur l'algorithme 5. Cette amélioration permettrait d'optimiser la prédiction de la valeur des actions à prendre selon l'état actuel. Ceci améliore donc la performance de l'algorithme [13].

### 6.3.3. Multi-Step Learning

Dans l'algorithme DQN classique, on a vu que la  $Q$ -fonction cible 6.1 s'écrivait, pour un état-action  $(s, a)$  d'une expérience :

$$r(s, a) + \gamma \max_{a_+} Q(s_+, a_+)$$

Dans le *Multi Step Learning* DQN, la valeur cible à l'instant  $T \in \mathbb{N}$  prend la forme :

$$\sum_{t=0}^{T-1} \gamma^t r(s_t, a_t) + \gamma^T \max_{a_+} Q(s_T, a_+)$$

Le code de cet algorithme est présenté dans l'algorithme 6. En faisant de la sorte, le choix de l'action

---

#### Algorithme 6 : Algorithme Multi Step Learning Deep $Q$ -Learning [14]

---

**Data :** On initialise les poids du modèle  $\theta$ , la mémoire  $\mathcal{M}$  de capacité  $M$ , l'environnement  $env$  et un agent  $agent$ .

**tant que** *On continue d'entraîner l'algorithme faire*

$env = reset()$ ;  $s_t, s_{t+1} = reset()$ ;  $episode = 0$ ;  $\varepsilon = 0$ ;  $\varepsilon_-$ ;

$game\_over = \mathbf{False}$ ;

**tant que**  $game\_over = \mathbf{False}$  **faire**

Avec probabilité  $\varepsilon$ , choisir  $a_t$  aléatoirement dans  $\mathcal{A}$ , sinon  $a_t \in \arg \max_{\mathcal{A}} Q_\theta(s_t, \cdot)$ ;

$agent$  effectue l'action  $a_t$  et récupère  $r_t$  et  $s_{t+1}$ , et l'état de  $game\_over$ ;

$R \leftarrow R + \gamma^t r_t$ ;

On ajoute  $(t, s_t, a_t, R, s_{t+1}, game\_over)$  à  $\mathcal{M}$ ;

On prélève un mini-batch  $\mathcal{B} = \{(t_i, s^i, a^i, R^i, s_+^i)_{i \leq b}\}$  d'expériences de  $\mathcal{M}$ ;

On entraîne  $Q_\theta$  sur  $\mathcal{B}$  et met à jour les poids : **for**  $i \leq b$  **do**

**si**  $game\_over$  **alors**

$y^i = R^i$

**sinon**

$y^i = R^i + \gamma^{t_i} \max_{a_+} Q_\theta(s^i, a_+)$

**fin**

**end**

On met à jour les poids  $\theta$  par descente de gradient avec la perte  $L$ ;

**fin**

On met à jour  $episode$ ,  $\varepsilon$ ,  $s_t$ ,  $s_{t+1}$

**fin**

---

optimale à un moment donné ne repose plus que sur les récompenses immédiates mais sur plusieurs étapes futures. Cela peut améliorer les performances du modèles dans certains environnement [14].

### 6.3.4. Distributional RL

Ensuite, le *Distributional Reinforcement Learning* DQN change du DQN classique du fait qu'il consiste à représenter la valeur d'une action comme une distribution plutôt qu'une seule valeur. Les actions à

effectuer ont ainsi un risque chacun à être prise, et l'agent peut apprendre la distribution des récompenses associées. Nous savons d'ores-et-déjà que la  $Q$ -fonction est définie par :

$$Q(s, a) = \mathbb{E}_\pi[G_t | S_t = s, A_t = a]$$

où  $G_t$  correspond au gain et est déterminée. Maintenant, nous décidons d'approcher ce gain par une distribution aléatoire  $d_t$  définie comme une somme de  $N \geq 1$  Dirac. Soit  $V_{\min}$  et  $V_{\max}$ , définis en amont, tel que le support de  $d_t$  se situe dans  $[V_{\min}; V_{\max}]$ , et  $\mathbf{z} = (z_i) \in \mathbb{R}^N$  avec  $z_i = V_{\min} + i \times \frac{V_{\max} - V_{\min}}{N - 1}$  les atomes de cette loi. La probabilité associée à chaque atome  $z_i$  de la loi est donnée par

$$p_\chi^i(s, a) = \frac{e^{\chi_i(s, a)}}{\sum_j e^{\chi_j(s, a)}}$$

où  $\chi : \mathcal{S} \times \mathcal{A} \mapsto \mathbb{R}^N$  est un modèle paramétrique à adapter. Nous notons cette distribution  $d_t = (\mathbf{z}, p_\chi(s_t, a_t))$ .

L'idée est donc d'adapter au mieux le modèle  $\chi$  pour approcher le plus précisément possible la vraie distribution des récompenses. Pour cela, on définit la distribution cible à partir de  $\bar{\chi}$  choisi :

$$\bar{d}_t = \left( r_{t+1} + \gamma^{t+1} \mathbf{z}, p_{\bar{\chi}} \left( s_t, \arg \max_{a \in \mathcal{A}} \mathbf{z}^T p_\chi(s_t, a) \right) \right)$$

La divergence de Kullbeck-Leibler permet enfin de mesurer la proximité de notre approximation avec cette distribution cible. Pour corriger  $\chi$  à chaque étape, on peut de nouveau utiliser un réseau de neurone.

Cette approche permettrait de mieux rendre compte de la variabilité des récompenses, et ainsi d'obtenir une meilleure estimation de l'action à prendre [15].

### 6.3.5. Noisy Nets

Nous avons précédemment introduit dans la partie 6.1.2 l'utilisation de stratégie  $\varepsilon$ -greedy dans les algorithmes DQN. Malheureusement, cette stratégie n'est pas forcément fiable pour certains environnements. Cela ne convient notamment pas pour des environnements de plus grande dimension.

L'idée est donc de conserver de la stochasticité dans l'algorithme, pour explorer potentiellement plus d'états, mais en utilisant des réseaux de neurones pour des couches de bruits plutôt. Le but est de rendre la sélection de l'action plus robuste. Les couches de bruits ajoutent une incertitude sur la prédiction de la valeur de la  $Q$ -fonction, ce qui empêche, comme avec la stratégie  $\varepsilon$ -greedy, de converger vers un maximum local.

Cette stratégie est implémentée dans l'algorithme 7 de la manière suivante : un réseau de neurones fournit un bruit qui va modifier l'estimation de la  $Q$ -fonction  $Q_\theta$ , de sorte que la  $Q$ -fonction estimée dépend à la fois du poids  $\theta$  du modèle et du bruit, nous noterons  $Q_\theta(s, a; \cdot)$ . Dans la partie 6.2.1, nous avons décrit les réseaux de neurones en les paramétrisant par une fonction  $f_\theta$  tel que, pour une entrée  $x$ , on récupère  $f_\theta(x)$  en sortie avec  $\theta$  le poids du réseau. Dans le cas d'un noisy network, le poids est défini par  $\mu + \Sigma \odot \varepsilon$  où  $\xi := (\mu, \Sigma)$  sont des paramètres à adapter, et  $\varepsilon$  est un bruit fixé, d'espérance nulle. Le symbole " $\odot$ " représente le produit élément par élément. Le poids du "noisy network" est donc mis à jour par descente de gradient également, mais en considérant la perte  $\bar{L} : \xi \mapsto \mathbb{E}_\varepsilon[L(\theta)]$ . De manière générale, si un réseau de neurone correspond à  $y = f_\theta(x) = \theta x + b$  avec  $b$  le biais, alors un noisy network de paramètres  $(\xi, \varepsilon)$  va modifier sa réponse en :

$$y = (\mu^\theta + \sigma^\theta \odot \varepsilon^\theta) x + \mu^b + \sigma^b \odot \varepsilon^b$$

où  $(\mu^\theta, \sigma^\theta) \in \mathbb{R}^{p \times q}$ ,  $(\mu^b, \sigma^b) \in \mathbb{R}^q$  sont des paramètres à apprendre et adapter, tandis que  $\varepsilon^\theta \in \mathbb{R}^{q \times p}$ , et  $\varepsilon^b \in \mathbb{R}^q$  sont des paramètres définis en amont [16]

**Algorithme 7 :** Algorithme Noisy DQN [16]

**Data :** On initialise les poids des modèles  $\theta$  et  $\xi$ , on choisit une variable aléatoire  $\varepsilon$ , la mémoire  $\mathcal{M}$  de capacité  $M$ , l'environnement  $env$  et un agent  $agent$ .

**tant que** *On continue d'entraîner l'algorithme faire*

$env = reset()$ ;  $s_t, s_{t+1} = reset()$ ;  $episode = 0$ ;  $\varepsilon = 0$ ;  $\varepsilon_-$ ;

$game\_over = \mathbf{False}$ ;

**tant que**  $game\_over = \mathbf{False}$ ;

**faire**

On simule  $\zeta \sim \varepsilon$ ;

On choisit  $a_t = \arg \max_{\mathcal{A}} Q_{\theta}(s, a; \xi, \zeta)$ ;

$agent$  effectue l'action  $a_t$  et récupère  $r_t$  et  $s_{t+1}$ , et l'état de  $game\_over$ ;

On ajoute  $(s_t, a_t, r_t, s_{t+1}, game\_over)$  à  $\mathcal{M}$ ;

On prélève un mini-batch  $\mathcal{B} = \{(s^i, a^i, r^i, s_+^i)\}_{i \leq b}$  d'expériences de  $\mathcal{M}$ ;

On entraîne le modèle sur  $\mathcal{B}$  et met à jour les poids : **for**  $i \leq b$  **do**

**si**  $game\_over$  **alors**

|  $y^i = r^i$

**sinon**

|  $y^i = r^i + \gamma \max_{a_+} Q_{\theta}(s_+^i, a_+)$

**fin**

**end**

On met à jour les poids  $\theta$  par descente de gradient avec la perte  $L$ , et tout les  $K$  épisodes celui de  $\xi$  avec  $\bar{L}$ ;

**fin**

On met à jour  $episode$ ,  $\varepsilon$ ,  $s_t$ ,  $s_{t+1}$

**fin**

### 6.3.6. Rainbow

L'algorithme Rainbow est finalement une combinaison de toutes ces extensions, conçu pour allier leurs avantages et en répondant à plusieurs problèmes de l'algorithme DQN classique. L'algorithme Rainbow utilise un réseau de neurones profond pour prédire la valeur  $Q$  pour chaque action dans un état donné. La priorité des transitions est alors déterminée en utilisant un algorithme de PER comme décrit ci-dessus pour sélectionner les transitions les plus importantes à utiliser pour la mise à jour du réseau de neurones. L'algorithme utilise également un Dueling Network pour une estimation plus précise des valeurs  $Q$ .

Les résultats présentés dans l'article [3] sur un grand nombre de jeux Atari montre que en combinant toutes ces extensions, l'algorithme Rainbow semble significativement plus efficace, rapide, et précis,

---

## Liste des références

---

David Silver. Lectures on reinforcement learning. URL : <https://www.davidsilver.uk/teaching/>, 2015.

Wilhelm Kirchgässner, Maximilian Schenke, Oliver Wallscheid, and Daniel Weber. Reinforcement learning course material. URL : [https://github.com/upb-lea/reinforcement\\_learning\\_course\\_materials](https://github.com/upb-lea/reinforcement_learning_course_materials), 2020. Paderborn University.

Matteo Hessel, Joseph Modayil, Hado van Hasselt, Tom Schaul, Georg Ostrovski, Will Dabney, Daniel Horgan, Bilal Piot, Mohammad Gheshlaghi Azar, and David Silver. Rainbow : Combining improvements in deep reinforcement learning. *CoRR*, abs/1710.02298, 2017.

Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym. *arXiv preprint arXiv :1606.01540*, 2016.

Andrea Lonza. *Hands-on reinforcement learning algorithms with Python : learn, understand, and develop smart algorithms for addressing AI challenges*. Packt Publishing, 1st edition. edition, 2019.

Nazia Habib. *Hands-on Q-learning with Python : practical Q-learning with OpenAI Gym, Keras, and TensorFlow*. Packt Publishing, 1st edition. edition, 2019.

Goldberg Yoav. *Neural network methods for natural language processing*. Synthesis lectures on human language technologies #37. San Rafael, California : Morgan and Claypool Publishers , C 2017, 2017.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin A. Riedmiller. Playing atari with deep reinforcement learning. *CoRR*, abs/1312.5602, 2013.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dhharshan Kumaran, Daan Wierstra, Shane Legg, and

Demis Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540) :529–533, February 2015.

Cong Feng, Mucun Sun, and Jie Zhang. Reinforced deterministic and probabilistic load forecasting via  $q$ -learning dynamic model selection. *IEEE Transactions on Smart Grid*, 11(2) :1377–1386, 2020.

Hado van Hasselt, Arthur Guez, and David Silver. Deep reinforcement learning with double q-learning. *CoRR*, abs/1509.06461, 2015.

Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. Prioritized experience replay, 2015.

Ziyu Wang, Tom Schaul, Matteo Hessel, Hado van Hasselt, Marc Lanctot, and Nando de Freitas. Dueling network architectures for deep reinforcement learning, 2015.

Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning : An Introduction*. The MIT Press, second edition, 2018.

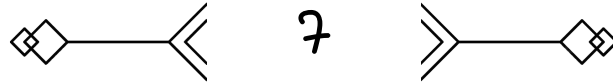
Marc G. Bellemare, Will Dabney, and Rémi Munos. A distributional perspective on reinforcement learning, 2017.

Meire Fortunato, Mohammad Gheshlaghi Azar, Bilal Piot, Jacob Menick, Ian Osband, Alex Graves, Vlad Mnih, Remi Munos, Demis Hassabis, Olivier Pietquin, Charles Blundell, and Shane Legg. Noisy networks for exploration, 2017.

## Troisième Partie : Approfondissement

,max size=0.9,widthh=

# Estimation de l'efficacité contact Tracing



Auriane GABAUT

Mars 2023

## Sommaire

|                                                                |            |
|----------------------------------------------------------------|------------|
| <b>Introduction et Motivations</b>                             | <b>131</b> |
| <b>7.1 Système Stochastique</b>                                | <b>132</b> |
| 7.1.1 Le modèle                                                | 132        |
| 7.1.2 Modèle Mathématique                                      | 133        |
| 7.1.3 Modélisation pour une grande population                  | 135        |
| <b>7.2 Approximate Bayesian Computation</b>                    | <b>140</b> |
| 7.2.1 Principe Général                                         | 140        |
| 7.2.2 Construction de l'ABC                                    | 140        |
| 7.2.3 Simulation                                               | 141        |
| <b>7.3 Estimation des paramètres pour des données simulées</b> | <b>142</b> |
| <b>Conclusion</b>                                              | <b>145</b> |

---

## Introduction et Motivations

---

L'épidémie de COVID-19 a montré la nécessité de comprendre les mécanismes de transmission d'une maladie infectieuse et d'évaluer l'efficacité des mesures de contrôle. Parmi ces mesures, le contact tracing est considéré comme une stratégie clé pour freiner la propagation de la maladie. Cependant, son efficacité dépend de plusieurs facteurs, tels que le temps de détection des contacts, la rapidité de la mise en place des mesures, leur bonne application, etc.

Le modèle SIR (*Susceptibles, Infectés, Rémis*) est un modèle mathématique couramment utilisé pour décrire la propagation des maladies infectieuses. Ce modèle divise la population en trois compartiments : les susceptibles, qui peuvent être infectés ; les infectieux, qui peuvent infecter d'autres personnes ; et les récupérés, qui ont guéri, sont décédés, ne transmettent plus, etc. Les paramètres clés du modèle SIR sont les différentes taux auquel chaque événement (détection, mort, infection...) s'effectue.

Dans ce rapport, nous allons examiner comment l'algorithme ABC (Approximate Bayesian Computation) peut être utilisé pour estimer les paramètres du modèle SIR dans le contexte de l'efficacité du contact tracing. L'algorithme ABC est une méthode statistique basée sur l'évaluation des probabilités conditionnelles pour des paramètres inconnus en fonction des données observées. Cette méthode est particulièrement utile lorsque le modèle mathématique est complexe et qu'il n'existe pas de solution analytique pour résoudre les équations. Nous nous basons sur le travail de Michael G.B. Blum et Viet Chi Tran [1].

---

## 7.1. Système Stochastique

---

### 7.1.1. Le modèle

Nous considérons un modèle SIR (*Susceptibles, Infectés, Rémis*) avec contact tracing. Nous considérons une population de taille  $N \in \mathbb{N}^*$  dans laquelle nous avons  $S : t \in \mathbb{R}_+ \mapsto S_t$  individus susceptibles d'être infectés, et  $I : t \in \mathbb{R}_+ \mapsto I_t$  individus infectés. Un individu susceptible devient infecté au taux individuel  $\lambda_1^N I_t$ , où  $\lambda_1$  est le taux d'infection dans une population de taille  $N$ .

Chaque individus infectés peut être détectés de deux manières différentes : par détection spontanée, ou par contact tracing. Les individus détectés jouent alors un rôle dans le contact tracing au cours du temps, dans le sens où ils permettent la détection d'autres individus qui leur sont proches. Nous définissons le poids d'un individu  $i \in \{1, \dots, N\}$  dans le processus de contact tracing comme  $t \geq T_i \mapsto e^{-c(t-T_i)}$  où  $T_i \in \mathbb{R}_+^*$  correspond au temps de détection du  $i^e$  individu. Ce poids est décroissant au cours du temps, chaque individu nouvellement infecté joue un rôle clefs dans le processus de détection mais l'effet diminue à mesure que le temps passe, puisqu'on peut supposer avoir tester tout ses proches par exemple. Autrement dit, chaque individu infecté est détecté spontanément au taux individuel  $\lambda_2^N$  ou détecté par contact tracing au taux individuel  $\lambda_3^N \sum_i e^{-c(t-T_i)}$ . Nous notons alors

$$R : t \in \mathbb{R}_+ \mapsto R_t = \sum_i e^{-c(t-T_i)}$$

la somme des individus détectés pondérés par leur poids dans le contact tracing.

Finalement, notre modèle contient les paramètres  $\theta^N = (\lambda_1^N, \lambda_2^N, \lambda_3^N, c)$ , et est représenté en figure 7.1. Ce modèle est un modèle simplifié de celui considéré dans l'article de Michael G.B. Blum et Viet Chi Tran [1], qui peut inclure des mouvements de populations, comme des naissances ou des décès.



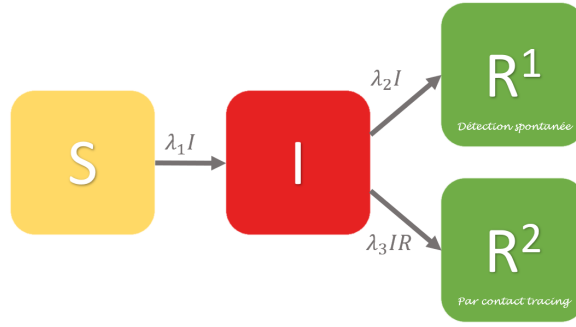


FIGURE 7.1 – Schéma du modèle SIR considéré.

### 7.1.2. Modèle Mathématique

Notre modèle décrit ci-dessus s'écrit :

$$\left\{ \begin{array}{l} S_t = S_0 - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s - I_s} Q^i(ds, d\theta) \\ I_t = I_0 + \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s - I_s} Q^i(ds, d\theta) - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_2^N I_s} Q^d(ds, d\theta) \\ \quad - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_3^N I_s - R_s} Q^c(ds, d\theta) \\ R_t = R_0 + \int_0^t \int \mathbb{1}_{\theta \leq \lambda_2^N I_s} Q^d(ds, d\theta) + \int_0^t \int \mathbb{1}_{\theta \leq \lambda_3^N I_s - R_s} Q^c(ds, d\theta) \\ \quad - c \int_0^t R_s ds \end{array} \right. \quad (7.1)$$

où les mesures  $Q^i$ ,  $Q^c$  et  $Q^d$  sont des mesures de poissos sur  $dsd\theta$  (les mesures de Lebesgues) sur  $\mathbb{R}_+^2$ . Ce système d'équation (7.1) peut être réécrit comme un système de semi-martingale :

#### Proposition 7.1.0 -

Nous pouvons écrire

$$\left\{ \begin{array}{l} S_t = S_0 - \int_0^t \lambda_1^N S_s I_s ds + M_t^S \\ I_t = I_0 + \int_0^t \left( \lambda_1^N S_s I_s - \lambda_2^N I_s - \lambda_3^N I_s R_s \right) ds + M_t^I \\ R_t = R_0 + \int_0^t \left( \lambda_2^N I_s + \lambda_3^N I_s R_s - c R_s \right) ds + M_t^R \end{array} \right. \quad (7.2)$$

où  $M^S$ ,  $M^I$ ,  $M^R$  sont des martingales de carrés intégrable, issue de 0 et de variation quadratique respective :

$$\int_0^t \lambda_1^N S_s I_s ds, \quad \int_0^t \lambda_1^N S_s I_s + \lambda_3^N I_s R_s ds, \quad \text{et} \quad \int_0^t \lambda_2^N I_s + \lambda_3^N I_s R_s ds$$

*Démonstration. Écriture en semi-martingale.* Soit la mesure de poisson compensée  $\tilde{Q}(ds, d\theta) =$

$Q(ds, d\theta) - dsd\theta$ , pour  $Q$  une mesure de poisson de mesure  $dsd\theta$  sur  $\mathbb{R}_+^2$ . Soit  $t \in \mathbb{R}_+$ , nous avons :

$$\begin{aligned} S_t &= S_0 - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s - I_s} Q^i(ds, d\theta) \\ &= S_0 - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s I_s} d\theta ds - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s - I_s} \tilde{Q}^i(ds, d\theta) \\ &= S_0 - \int_0^t \lambda_1^N S_s I_s ds - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s - I_s} \tilde{Q}^i(ds, d\theta) \end{aligned}$$

Nous posons  $M_t^S = - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s - I_s} \tilde{Q}^i(ds, d\theta)$ .  $M_t^S$  vérifie pour  $s < t$  :

$$\begin{aligned} \mathbb{E}[M_t^S | \mathcal{F}_s] &= -\mathbb{E} \left[ \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_u - I_u} \tilde{Q}^i(du, d\theta) | \mathcal{F}_s \right] \\ &= -\mathbb{E} \left[ \int_0^s \int \mathbb{1}_{\theta \leq \lambda_1^N S_u - I_u} \tilde{Q}^i(du, d\theta) | \mathcal{F}_s \right] - \mathbb{E} \left[ \int_s^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_u - I_u} \tilde{Q}^i(du, d\theta) | \mathcal{F}_s \right] \\ &= -\int_0^s \int \mathbb{1}_{\theta \leq \lambda_1^N S_u - I_u} \tilde{Q}^i(du, d\theta) - \mathbb{E} \left[ \int_s^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_u - I_u} \tilde{Q}^i(du, d\theta) \right] \end{aligned}$$

et

$$\mathbb{E} \left[ \int_s^t \int |\mathbb{1}_{\theta \leq \lambda_1^N S_u - I_u}| dud\theta \right] < (t-s)\lambda_1^N N^2 < \infty$$

donc :

$$\mathbb{E}[M_t^S | \mathcal{F}_s] = - \int_0^s \int \mathbb{1}_{\theta \leq \lambda_1^N S_u - I_u} \tilde{Q}^i(du, d\theta) = M_s^S$$

$M^S$  est donc bien une martingale, issue de 0 puisque  $M_0^S = 0$ .

De la même manière, nous avons :

$$I_t = I_0 + \int_0^t \left( \lambda_1^N S_s I_s - \lambda_2^N I_s - \lambda_3^N I_s R_s \right) + M_t^I$$

avec la martingale

$$M_t^I = \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s - I_s} \tilde{Q}^i(ds, d\theta) - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_2^N I_s} \tilde{Q}^d(ds, d\theta) - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_3^N I_s - R_s} \tilde{Q}^c(ds, d\theta)$$

et

$$R_t = R_0 + \int_0^t \left( \lambda_2^N I_s + \lambda_3^N I_s R_s - c R_s \right) ds + M_t^R$$

où  $M_t^R$  est la martingale telle que :

$$M_t^R = \int_0^t \int \mathbb{1}_{\theta \leq \lambda_2^N I_s} \tilde{Q}^d(ds, d\theta) + \int_0^t \int \mathbb{1}_{\theta \leq \lambda_3^N I_s - R_s} \tilde{Q}^c(ds, d\theta)$$

*Variation Quadratique.* Soit  $t \in \mathbb{R}_+$ , on note :

$$\begin{cases} M_{1,t}^I &= \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s - I_s} \tilde{Q}^i(ds, d\theta) \\ M_{2,t}^I &= - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_2^N I_s} \tilde{Q}^d(ds, d\theta) \\ M_{3,t}^I &= - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_3^N I_s - R_s} \tilde{Q}^c(ds, d\theta) \end{cases}$$

De sorte que  $M_t^I = M_{1,t}^I + M_{2,t}^I + M_{3,t}^I$ . Nous avons donc :

$$\begin{aligned} \mathbb{E}[\langle M^I \rangle_t] &= \mathbb{E}[(M_{1,t}^I + M_{2,t}^I + M_{3,t}^I)^2] \\ &= \mathbb{E}[(M_{1,t}^I)^2] + \mathbb{E}[(M_{2,t}^I)^2] + \mathbb{E}[(M_{3,t}^I)^2] \\ &\quad + 2 \left( \mathbb{E}[M_{1,t}^I M_{2,t}^I] + \mathbb{E}[M_{1,t}^I M_{3,t}^I] + \mathbb{E}[M_{2,t}^I M_{3,t}^I] \right) \end{aligned}$$

et, comme  $\mathbb{E} \left[ \int_0^t \int |\mathbb{1}_{\theta \leq \lambda_1^N S_u I_u}| ds d\theta \right] \leq \lambda_1^N N^2 t < +\infty$  :

$$\begin{aligned} \mathbb{E}[(M_{1,t}^I)^2] &= \mathbb{E} \left[ \left( \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s I_s} \tilde{Q}(ds, d\theta) \right)^2 \right] \\ &= \mathbb{E} \left[ \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1^N S_s I_s} ds d\theta \right] \\ &= \mathbb{E} \left[ \int_0^t \lambda_1^N S_s I_s ds \right] \end{aligned}$$

Donc :

$$\begin{cases} \mathbb{E}[(M_{1,t}^I)^2] &= \mathbb{E} \left[ \int_0^t \lambda_1^N S_s I_s ds \right] \\ \mathbb{E}[(M_{2,t}^I)^2] &= \mathbb{E} \left[ \int_0^t \lambda_2^N I_s ds \right] \\ \mathbb{E}[(M_{3,t}^I)^2] &= \mathbb{E} \left[ \int_0^t \lambda_3^N I_s R_s ds \right] \end{cases}$$

Montrons maintenant que  $\mathbb{E}[M_{1,t}^I M_{2,t}^I] = \mathbb{E}[M_{1,t}^I M_{3,t}^I] = \mathbb{E}[M_{2,t}^I M_{3,t}^I] = 0$ . Pour cela, nous appliquons la formule d'Itô à  $M_{i,t}^I$  et  $M_{j,t}^I$  avec  $\varphi : (x, y) \mapsto xy$ , d'où :

$$\begin{aligned} M_{1,t}^I M_{2,t}^I &= \int_0^t \int \left[ (M_{1,s^-}^I + \mathbb{1}_{\theta \leq \lambda_1^N S_s I_s} - M_{1,s^-}^I) M_{2,s^-}^I Q^i(ds, d\theta) \right. \\ &\quad \left. - \int_0^t \int M_{2,s^-}^I \times \mathbb{1}_{\theta \leq \lambda_1^N S_s I_s} ds d\theta - \int_0^1 \int M_{1,t}^I \times \mathbb{1}_{\theta \leq \lambda_2^N I_s} Q^d(ds, d\theta) \right. \\ &\quad \left. + \int_0^t \int M_{1,s^-}^I \times \mathbb{1}_{\theta \leq \lambda_2^N I_s} ds d\theta \right] \end{aligned}$$

Donc  $\mathbb{E}[M_{1,t}^I M_{2,t}^I] = 0$  et de la même manière

$$\mathbb{E}[M_{1,t}^I M_{2,t}^I] = \mathbb{E}[M_{1,t}^I M_{3,t}^I] = \mathbb{E}[M_{2,t}^I M_{3,t}^I] = 0$$

Finalement,

$$\mathbb{E}[\langle M^I \rangle_t] = \mathbb{E} \left[ \int_0^t \left( \lambda_1^N S_s I_s + \lambda_2^N I_s + \lambda_3^N I_s R_s \right) ds \right]$$

Nous procédons de la même manière pour  $M^S$  et  $M^R$ , et finalement :

$$\begin{cases} \mathbb{E}[\langle M^S \rangle_t] &= \mathbb{E} \left[ \int_0^t \left( \lambda_1^N S_s I_s \right) ds \right] \\ \mathbb{E}[\langle M^I \rangle_t] &= \mathbb{E} \left[ \int_0^t \left( \lambda_1^N S_s I_s + \lambda_2^N I_s + \lambda_3^N I_s R_s \right) ds \right] \\ \mathbb{E}[\langle M^R \rangle_t] &= \mathbb{E} \left[ \int_0^t \left( \lambda_2^N I_s + \lambda_3^N I_s R_s \right) ds \right] \end{cases}$$

□

### 7.1.3. Modélisation pour une grande population

#### Reformulation du système

Dans le but de trouver la dynamique du processus en grande population, donc pour  $N$  asymptotiquement grand, nous souhaitons faire tendre  $N$  vers l'infini dans le système 7.1. Pour cela, nous formulons les hypothèses suivantes :

- i. Dans le cas d'une population de taille  $N$ , nos paramètres  $(\lambda_1^N, \lambda_2^N, \lambda_3^N)$  sont de la forme  $(\lambda_1/N, \lambda_2, \lambda_3/N)$ , où  $(\lambda_1, \lambda_2, \lambda_3)$  sont les paramètres généraux de l'épidémie. Le paramètre  $\lambda_2^N$  est supposé égale au paramètre général  $\lambda_2$ , puisqu'il n'est pas supposé dépendre de la population mais de l'individu infecté seulement.

- ii.  $\lim_{N \rightarrow +\infty} I_t = \varepsilon > 0$

Soit  $(S_t^N, I_t^N, R_t^N) = \left(\frac{S_t}{N}, \frac{I_t}{N}, \frac{R_t}{N}\right)$  les taux d'individus susceptibles, infectés et détectés dans le cas d'une population de taille  $N$ . Nous avons alors :

$$\begin{cases} S_t^N &= S_0^N - \frac{1}{N} \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1 S_s^N I_s^N} Q^i(ds, d\theta) \\ I_t^N &= I_0^N + \frac{1}{N} \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1 S_s^N I_s^N} Q^i(ds, d\theta) - \frac{1}{N} \int_0^t \int \mathbb{1}_{\theta \leq \lambda_2 I_s^N} Q^d(ds, d\theta) \\ &\quad - \frac{1}{N} \int_0^t \int \mathbb{1}_{\theta \leq \lambda_3 I_s^N R_s^N} Q^c(ds, d\theta) \\ R_t^N &= R_0^N + \frac{1}{N} \int_0^t \int \mathbb{1}_{\theta \leq \lambda_2 I_s^N} Q^d(ds, d\theta) + \frac{1}{N} \int_0^t \int \mathbb{1}_{\theta \leq \lambda_3 I_s^N R_s^N} Q^c(ds, d\theta) \\ &\quad - c \int_0^t R_s^N ds \end{cases} \quad (7.3)$$

Nous pouvons de nouveau réécrire le système en utilisant la proposition 1 :

$$\begin{cases} S_t^N &= S_0^N - \int_0^t \lambda_1 S_s^N I_s^N ds + M_t^{N,S} \\ I_t^N &= I_0^N + \int_0^t \left( \lambda_1 S_s^N I_s^N - \lambda_2 I_s^N - \lambda_3 I_s^N R_s^N \right) ds + M_t^{N,I} \\ R_t^N &= R_0^N + \int_0^t \left( \lambda_2 I_s^N + \lambda_3 I_s^N R_s^N - c R_s^N \right) ds + M_t^{N,R} \end{cases} \quad (7.4)$$

avec les martingales

$$\begin{cases} M_t^{N,S} &= \frac{1}{N} \times \left( \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1 S_s^N I_s^N} \tilde{Q}^i(ds, d\theta) \right) \\ M_t^{N,I} &= \frac{1}{N} \times \left( \int_0^t \int \mathbb{1}_{\theta \leq \lambda_1 S_s^N I_s^N} \tilde{Q}^i(ds, d\theta) - \int_0^t \int \mathbb{1}_{\theta \leq \lambda_2 I_s^N} \tilde{Q}^d(ds, d\theta) \right) \\ M_t^{N,R} &= \frac{1}{N} \times \left( \int_0^t \int \mathbb{1}_{\theta \leq \lambda_2 I_s^N} \tilde{Q}^d(ds, d\theta) + \int_0^t \int \mathbb{1}_{\theta \leq \lambda_3 I_s^N R_s^N} \tilde{Q}^c(ds, d\theta) \right) \end{cases}$$

de variation quadratique respective :

$$\frac{1}{N} \int_0^t \lambda_1 S_s^N I_s^N ds, \quad \frac{1}{N} \int_0^t \lambda_1 S_s^N I_s^N + \lambda_3 I_s^N R_s^N ds, \quad \text{et} \quad \frac{1}{N} \int_0^t \lambda_2 I_s^N + \lambda_3 I_s^N R_s^N ds$$

Existence d'une limite

Définition 7.1.1 -

Une famille de probabilités  $\mathcal{P}$  sur un espace métrique  $(E, d)$  est tendue si

$$\forall \varepsilon > 0, \exists K_\varepsilon \text{ compact } \subset E, \forall P \in \mathcal{P}, P(K_\varepsilon^C) < \varepsilon$$

Théorème 7.1.1 - Théorème de Prokhorov

Si  $(E, d)$  est séparable et complet, Alors

$$\mathcal{P} \text{ tendue } \Leftrightarrow \mathcal{P} \text{ relativement compact.}$$

Les fonctions  $S^N, I^N, R^N$  sont des fonctions de l'espace de Skorohod  $\mathbb{D}([0; T])$ , c'est-à-dire que  $S^N, I^N, R^N$  sont continues à droite et limitées à gauche sur  $[0; T]$ . Muni de la distance de Prokhorov

$$d^\circ(y, z) = \inf_{\lambda \in \Lambda} \left\{ \max \left( \|y \circ \lambda - z\|_\infty, \sup_{s \leq t} \left| \log \left( \frac{\lambda(t) - \lambda(s)}{t - s} \right) \right| \right) \right\}$$

où  $\Lambda = \{\lambda : [0; 1] \mapsto [0; 1] \text{ bijective avec } \lambda(0) = 0, \lambda(1) = 1\}$ , l'espace  $\mathbb{D}([0; 1])$  est séparable et complet.

Si nous montrons que la famille des distributions de  $X^N = (S^N, I^N, R^N)$  est tendue sur  $(\mathbb{D}, d^\circ)$ , alors, par le théorème 1 de Prokhorov, la famille est également relativement compacte. La famille de loi  $(\mathcal{L}(X^N))$  admet alors une valeur d'adhérence dans l'espace des lois  $\mathcal{P}(\mathbb{D})$ . Nous obtenons alors que, si la valeur d'adhérence est unique, alors  $X^N$  converge en loi vers la valeur d'adhérence.

Théorème 7.1.2 -

Soit  $(X^N)$  une suite de semi-martingale  $X^N = A^N + M^N$ , pour tout  $N \in \mathbb{N}^*$ . Si :

- $[T_1] \forall t \in [0 : T], \sup_N \mathbb{E}[|A_t^N|] < +\infty$  et  $\sup_N \mathbb{E}[< M^N >_t] < +\infty$ ,
- $[A] \forall \varepsilon > 0, \forall \eta > 0, \exists \delta > 0, N_0 \in \mathbb{N}, \forall N \geq N_0, \forall (\sigma_N, \tau_N)$  temps d'arrêt tel que  $\sigma_N \leq \tau_N < \sigma_N + \delta$ , alors
 
$$\sup_{N \geq N_0} \mathbb{P}(|A_{\tau_N}^N - A_{\sigma_N}^N| \geq \zeta) \leq \varepsilon \text{ et },$$

$$\sup_{N \geq N_0} \mathbb{P}(|< M^N >_{\tau_N} - < M^N >_{\sigma_N}| > \eta) < \varepsilon$$

Alors la famille des lois de  $(X_N)_{N \in \mathbb{N}^*}$  est tendue.

Nous pouvons appliquer le théorème 2 à nos solutions de (7.4).

$[T_1]$ . Nous avons toujours  $S_t^N, I_t^N, R_t^N \in [0; 1]$  pour tout temps  $t \in [0; T]$ . Donc, pour

$$S_t^N = S_0^N - \underbrace{\int_0^t \lambda_1 S_s^N I_s^N ds}_{=A_t^{N,S}} + M_t^{N,S}$$

:

$$\begin{aligned} \mathbb{E}[|A_t^{N,S}|] &\leq \mathbb{E}\left[\underbrace{S_0^N}_{\leq 1} + \int_0^t \underbrace{\lambda_1 S_s^N I_s^N}_{\leq 1} ds\right] \\ &\leq \mathbb{E}\left[1 + \lambda_1 \underbrace{t}_{\leq T}\right] \\ &\leq 1 + \lambda_1 T \end{aligned}$$

D'où  $\sup_{t \in [0; T]} \mathbb{E}[|A_t^{N,S}|] \leq 1 + \lambda_1 T$ . Nous avons de plus

$$\begin{aligned} \mathbb{E}[< M^{N,S} >_t] &= \mathbb{E}\left[\left(M_t^{N,S}\right)^2\right] \\ &= \mathbb{E}\left[\frac{1}{N^2} \times \left(\int_0^1 \int \mathbb{1}_{\theta \leq \lambda_1 S_s^N I_s^N} \tilde{Q}^i(ds, d\theta)\right)^2\right] \\ &= \frac{1}{N^2} \mathbb{E}\left[\int_0^t \int \mathbb{1}_{\lambda_1 S_s^N I_s^N} ds d\theta\right] \\ &= \frac{1}{N} \mathbb{E}\left[\int_0^t \lambda_1 S_s^N I_s^N ds\right] \\ &\leq \frac{1}{N} \lambda_1 T. \end{aligned}$$

De la même manière, on a

$$\sup_{t \in [0; T]} \mathbb{E}[< M^{N,S} >_t] \leq \frac{1}{N} \lambda_1 T.$$

Le critère  $[T_1]$  est donc vérifié pour  $I^N$  sur  $[0; T]$ , et nous procédons exactement de la même manière pour  $S^N, R^N$ .

[A]. Soit  $\varepsilon > 0, \eta > 0$ , on cherche à trouver  $\delta > 0$  tel que pour tout  $N \geq N_0, \forall (\sigma_N, \tau_N)$  temps d'arrêt tel que  $\sigma_N \leq \tau_N < \sigma_N + \delta$

$$\begin{aligned} - \sup_{N \geq N_0} \mathbb{P}(|A_{\tau_N}^{N,S} - A_{\sigma_N}^{N,S}| > \eta) &\leq \varepsilon \\ - \sup_{N \geq N_0} \mathbb{P}(|\langle M^{N,S} \rangle_{\tau_N} - \langle M^{N,S} \rangle_{\sigma_N}| > \eta) &< \varepsilon \end{aligned}$$

Pour cela, remarquons que

$$\begin{aligned} \mathbb{P}(|A_{\tau_N}^{N,S} - A_{\sigma_N}^{N,S}| > \eta) &\leq \frac{1}{\eta} \mathbb{E}[|A_{\tau_N}^{N,S} - A_{\sigma_N}^{N,S}|] && \text{par Markov} \\ &\leq \frac{1}{\eta} \mathbb{E} \left[ \left| \int_{\sigma_N}^{\tau_N} (\lambda_1 S_s^n I_s^N ds) \right| \right] \\ &\leq \frac{1}{\eta} \mathbb{E} \left[ \int_{\sigma_N}^{\tau_N} \lambda_1 ds \right] \\ &\leq \frac{\lambda_1 \delta}{\eta} && \text{car } (\tau_N - \sigma_N) < \delta. \end{aligned}$$

Donc si  $\delta \leq \frac{\varepsilon \eta}{\lambda_1} = \delta_0$ , on a bien

$$P(|A_{\tau_N}^{N,S} - A_{\sigma_N}^{N,S}| > \eta) \leq \varepsilon, \quad \forall N \geq N_0.$$

De la même manière, pour  $N \geq N_0$ ,

$$\begin{aligned} \mathbb{P}(|\langle M^{N,S} \rangle_{\tau_N} - \langle M^{N,S} \rangle_{\sigma_N}| > \eta) &\leq \frac{1}{\eta} \mathbb{E} [|\langle M^{N,S} \rangle_{\tau_N} - \langle M^{N,S} \rangle_{\sigma_N}|] \\ &\leq \frac{1}{\eta N} \mathbb{E} \left[ \left| \int_{\sigma_N}^{\tau_N} \int_0^t \lambda_1 S_s^n I_s^N ds \right| \right] \\ &\leq \frac{1}{\eta N_0} \lambda_1 \delta. \end{aligned}$$

Donc si  $\delta \leq \frac{\varepsilon \eta N_0}{\lambda_1} = \delta_1$ , on a bien

$$\sup_{N \geq N_0} \mathbb{P}(|\langle M^N \rangle_{\tau_N} - \langle M^N \rangle_{\sigma_N}| > \eta) < \varepsilon.$$

Les deux égalités sont donc vérifiées simultanément pour  $\delta \leq \min(\delta_0, \delta_1)$ . En procédant de la même manière, nous trouvons les mêmes résultat pour  $I^N$  et  $R^N$ .

Par le théorème 2, la famille des lois de  $(S_t^N, I_t^N, R_t^N)$  est relativement compact dans l'ensemble des lois de  $\mathbb{D}$ . Elle admet donc une valeur d'adhérence  $(\bar{S}, \bar{I}, \bar{R})$  telle que  $(S^{\varphi(N)}, I^{\varphi(N)}, R^{\varphi(N)}) \xrightarrow{\mathcal{L}} (\bar{S}, \bar{I}, \bar{R})$ . D'après le théorème 3, nous pouvons nous ramener au cas où  $(S^N, I^N, R^N)_{N \in \mathbb{N}}$  converge presque sûrement vers  $(\bar{S}, \bar{I}, \bar{R})$

### Théorème 7.1.3 - Théorème de Skorohod

Soit une suite de variables aléatoires  $(X_n)_{n \in \mathbb{N}}$  convergeant en loi vers une variable aléatoire  $X$ . Alors on peut construire  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  et  $(\tilde{X}_n)_{n \in \mathbb{N}}, \tilde{X}$ , sur  $\tilde{\Omega}$  tels que :

- i.  $\forall N \in \mathbb{N}, X_N \stackrel{d}{=} \tilde{X}_N$ ,
- ii.  $\forall N \in \mathbb{N}, X \stackrel{d}{=} \tilde{X}$
- iii.  $\tilde{X}_N \xrightarrow[N \rightarrow \infty]{p.s.} \tilde{X}$

Unicité de la limite p.s.

Soit la fonction

$$\begin{aligned}\psi_t(s, i, r) &= \left| s_t - s_0 + \int_0^t \lambda_1 s_s i_s ds \right| \\ &\quad + \left| i_t - i_0 - \int_0^t (\lambda_1 s_s i_s - \lambda_2 i_s - \lambda_3 i_s r_s) ds \right| \\ &\quad + \left| r_t - r_0 - \int_0^t (\lambda_2 i_s + \lambda_3 i_s r_s - c r_s) ds \right|\end{aligned}\tag{7.5}$$

et donc, pour  $(S^N, I^N, R^N)$ , nous avons  $\psi_t(S^N, I^N, R^N) = |M_t^{N,S}| + |M_t^{N,I}| + |M_t^{N,R}|$ .

Proposition 7.1.3 -

$\psi$  vérifie

$$\lim_{N \rightarrow +\infty} \psi_t(S^N, I^N, R^N) = 0, \text{ p.s.}$$

et  $\psi : t \in [0; T] : \psi_t$  n'est pas continue sur un nombre au plus dénombrable de points.

*Démonstration. Limite en l'infini.* Nous avons

$$\begin{aligned}\mathbb{E} [\psi_t^2(S^N, I^N, R^N)] &\leq \mathbb{E}[|M_t^{N,S}|^2] + \mathbb{E}[|M_t^{N,I}|^2] + \mathbb{E}[|M_t^{N,R}|^2] && \text{car } \mathbb{E}[M_t^{N,i} M_t^{N,j}] = 0 \\ &\leq 2 \frac{1}{N} (\lambda_1 + \lambda_2 + \lambda_3) && \text{car } S^N, I^N, R^N \in [0; 1]\end{aligned}$$

*Continuité de  $\psi_t$ .* Soit  $\varphi \in C^0(\mathbb{D})$ , nous avons les fonctions

$$y \in \mathbb{D} \mapsto \int_0^t y_s ds, \quad y \in \mathbb{D} \mapsto \varphi(y), \quad \text{et} \quad y \in \mathbb{D} \mapsto \sup_{t \leq T} y_t$$

continues sur  $\mathbb{D}([0; T])$ . En revanche, la fonction  $y \mapsto y_{t_0}$  pour  $t_0 \in [0; T]$  n'est pas continue un en nombre au plus dénombrable de points, correspondant au temps de saut de la fonction continue à droite, limitée à gauche  $y \in \mathbb{D}([0; T])$ .  $\square$

Nous notons l'ensemble  $\mathcal{T}$  l'ensemble de ses points de discontinuité, nous avons alors pour tout  $t \notin \mathcal{T}$  où  $\psi$  est continue :

$$\psi_t(S^N, I^N, R^N) \xrightarrow{N \rightarrow +\infty} \psi(\bar{S}, \bar{I}, \bar{R})$$

et nous avons alors  $\psi_t(\bar{S}, \bar{I}, \bar{R}) = 0$ , d'où :

$$\begin{cases} \bar{S}_t &= \bar{S}_0 - \int_0^t \lambda_1 \bar{S}_s \bar{I}_s ds \\ \bar{I}_t &= \bar{I}_0 + \int_0^t -\lambda_1 \bar{S}_s \bar{I}_s - \lambda_2 \bar{I}_s - \lambda_3 \bar{I}_s \bar{R}_s ds \\ \bar{R}_t &= \bar{R}_0 + \int_0^t \lambda_2 \bar{I}_s + \lambda_3 \bar{I}_s \bar{R}_s - c \bar{R}_s ds \end{cases}\tag{7.6}$$

Toutes valeurs d'adhérence de  $(S^N, I^N, R^N)$  vérifient l'équation 7.6, donc on a unicité des solutions, elle est de plus continue.

---

## 7.2. Approximate Bayesian Computation

---

### 7.2.1. Principe Général

L'algorithme ABC, ou Approximate Bayesian Computation, est une méthode statistique qui permet d'estimer les paramètres inconnus d'un modèle à partir de données observées. Cette méthode est particulièrement utile lorsque le modèle est complexe et que la vraisemblance est difficile à calculer. L'idée centrale de l'algorithme ABC est de simuler des données à partir du modèle en utilisant différents paramètres, puis de comparer ces données simulées aux données observées en utilisant une mesure de distance. Si la distance est suffisamment petite, alors les paramètres correspondants sont considérés comme plausibles. L'algorithme ABC est souvent utilisé en biologie, en écologie, en économie et en physique, où les données peuvent être rares ou bruyantes, et où les modèles sont souvent complexes et difficilement estimables de manière analytique.

Pour appliquer l'algorithme ABC à notre modèle (7.3), nous procédons de la manière suivante :

- i) Nous définissons une loi à priori sur les paramètres  $\theta \sim \pi(\theta)$ .
- ii) Nous simulons  $N \geq 1$  simulations  $(s_i)_{i \leq N}$  de paramètres  $(\theta_i)_{i \leq N} \stackrel{iid}{\sim} \pi(\theta)$ .
- iii) Pour chaque simulations  $i \leq N$ , nous calculons la distance  $d_i$  de la simulation aux données observées. Nous acceptons la simulation si  $d_i \leq \delta$ , où  $\delta$  est le quantile d'ordre 0.1 de la loi empirique des  $(d_i)$ . Chaque simulation acceptée a alors un poids  $w_i$  d'assigner, définit par

$$w_i = K(d_i)$$

où  $K$  est un noyau statistique, c'est-à-dire une fonction réelle, symétrique, intégrable de norme  $\|K\|_1 = 1$ . Le noyau  $K$  admet son maximum en 0.

- iv) Finalement, la loi a posteriori est estimée par :

$$\frac{\sum_{i \in I} w_i \delta_{\theta_i}}{\sum_{i \in I} w_i}$$

où  $\delta_\theta$  est la mesure de Dirac en  $\theta$ , et  $I \subset \{1, \dots, N\}$  est le sous-ensemble des simulations  $i \leq N$  acceptées.

### 7.2.2. Construction de l'ABC

Nous supposons dans un premier temps pouvoir simuler une trajectoire du modèle (7.3) à partir de paramètres  $\theta \sim \pi(\theta)$ , tirés à partir d'une loi prior  $\pi$  choisit. Nous souhaitons particulièrement estimer l'efficacité du contact tracing, donc de  $c$ . Pour cela, on considère les fonctions  $R^1$  et  $R^2$  le nombre d'individu detecté par détection spontanée, et par contact tracing respectivement. Leur somme est donc



le nombre total d'individus détectés. On pose pour  $l \in \{1, 2\}$

$$d_i = \|R_{obs}^l - R_i^l\|_1 = \int_0^T |R_{obs,s}^l - R_{i,s}^l| ds, \quad \forall i \in \{1, \dots, N\}$$

où  $R_i^1, R_i^2$  sont les fonctions  $R^1, R^2$  de la  $i^e$  simulation. Le poids d'une simulation est alors

$$w_i = K_{\delta_1}(\|R_{obs}^1 - R_i^1\|) K_{\delta_2}(\|R_{obs}^2 - R_i^2\|), \quad \forall i \in \{1, \dots, N\}$$

où  $K_\delta$  est le noyau d'Epanechnikov défini par :

$$K_\delta(t) = \begin{cases} \frac{4}{3\delta} \left(1 - \frac{t^2}{\delta^2}\right), & |t| \leq \delta \\ 0, & |t| > \delta \end{cases}$$

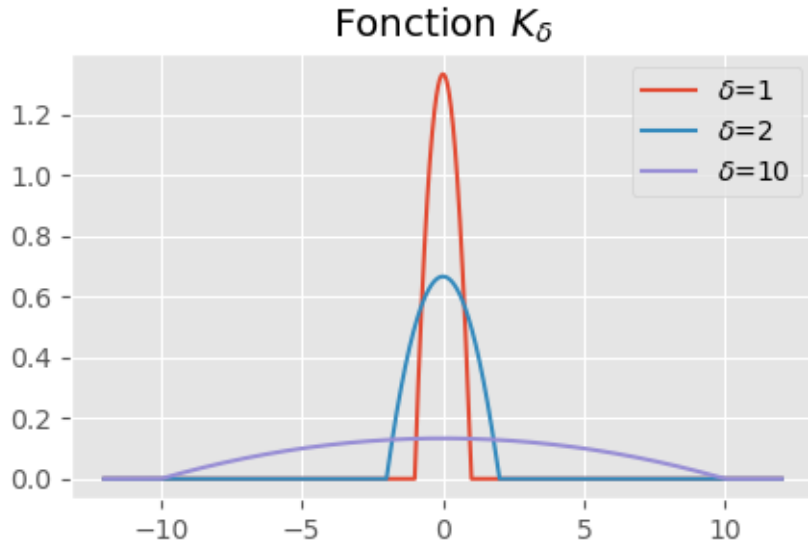


FIGURE 7.2 – Graph du noyau d'Epanechnikov, pour  $\delta = 1, 2$  et  $10$ .

Les paramètres  $\delta_1, \delta_2$  sont choisis de la façon suivante. Nous fixons  $\delta_j$  comme le quantile  $P_{\delta_j}$  de la distribution empirique calculé  $\|R_i - R_{obs}\|$ . On choisit  $P_{\delta_1} = P_{\delta_2} = 0.01$  de sorte que 1% des simulations  $(R_i)_{i \leq N}$  les plus proches de  $R_{obs}$  ont un poids non nul.

### 7.2.3. Simulation

Notre modèle décrit un processus de Markov de saut. Ici, les processus considérés sont des processus de Poisson sur  $\mathbb{R}_+$ . Nous rappelons les propriétés suivantes pour les processus de sauts purs.

#### Proposition 7.2.0 -

Nous notons  $T_i$  le temps du  $i^e$  saut, et la convention  $T_0 = 0$ .

La loi des temps d'attentes  $(T_i - T_{i-1})_{i \geq 1}$  est une loi exponentielle de paramètre  $\Lambda_{i-}$ , correspondant à la somme des taux des  $n$  différents sauts avant le saut. De plus, le saut a alors une probabilité  $\frac{\lambda_{i-}^n}{\Lambda_{i-}}$  d'être le  $n^e$  saut, où  $\lambda_{i-}^n$  est le taux du  $n^e$  saut avant le  $i^e$  saut.

Nous procédons donc de la manière suivante pour simuler le processus décrit par notre système, en approximant le processus  $R$  à un processus de saut pur, et donc de taux constant entre chaque saut. Nous supposons construit  $S_t, I_t, R_t$  pour un certain temps  $t$ .

1. On simule  $e \sim \mathcal{E}(\lambda_1 S_t I_t + \lambda_2 I_t + \lambda_3 I_t R_t)$ , le temps d'attente jusqu'au prochain temps de saut  $t_+ = t + e$ .
2. On calcule la valeur de  $S, I$  et  $R$  avant le prochain temps de saut :

$$\begin{cases} S_{t_+} &= S_t \\ I_{t_+} &= I_t \\ R_{t_+} &= \sum_i e^{-c(t_+ - T_i)} = R_t \times e^{-ce} \end{cases}$$

3. On simule  $u \sim \mathcal{U}([0 ; \lambda_1 S_{t_+} I_{t_+} + \lambda_2 I_{t_+} + \lambda_3 I_{t_+} R_{t_+}])$ , puis :
  - Si  $0 \leq u < \lambda_1 S_{t_+} I_{t_+}$ , alors l'évènement est l'infection d'un individu susceptible.
  - Si  $\lambda_1 S_{t_+} I_{t_+} \leq u < \lambda_1 S_{t_+} I_{t_+} + \lambda_2 I_{t_+}$ , alors l'évènement est la détection spontanée d'un individu infecté.
  - Si  $\lambda_1 S_{t_+} I_{t_+} + \lambda_2 I_{t_+} \leq u < \lambda_1 S_{t_+} I_{t_+} + \lambda_2 I_{t_+} + \lambda_3 I_{t_+} R_{t_+}$ , alors l'évènement est la détection par contact tracing d'un individu infecté.

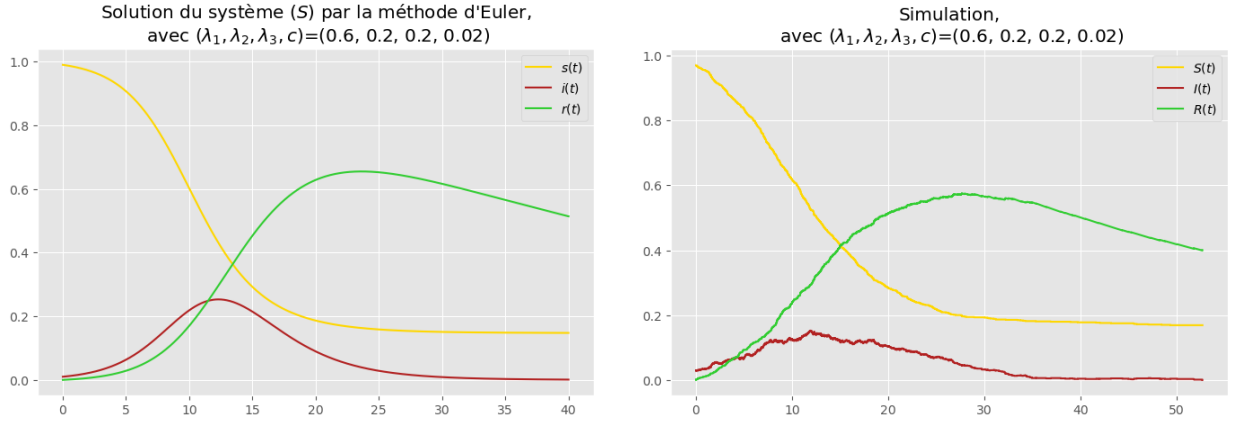


FIGURE 7.3 – Système limite et simulation pour les paramètres  $\lambda_1 = 0.6, \lambda_2 = 0.2, \lambda_3 = 0.2, c = 0.02$ .

Les codes utilisés sont disponibles sur [https://github.com/aurianegbt/HIV\\_stats](https://github.com/aurianegbt/HIV_stats).

---

### 7.3. Estimation des paramètres pour des données simulées

---

Pour tester l'ABC construit, disponible sur [https://github.com/aurianegbt/HIV\\_stats](https://github.com/aurianegbt/HIV_stats). On génère le processus pour

$$\lambda_1 = 0.6, \lambda_2 = 0.2, \lambda_3 = 0.2, c = 0.01.$$

Les données considérées sont représentées sur la figure 7.4.

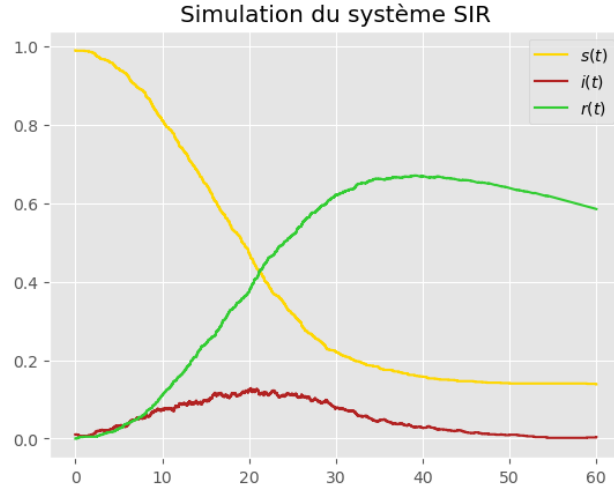


FIGURE 7.4 – Données observées.

Nous utilisons l'algorithme ABC pour  $N = 40000$  simulations, avec les prior respective  $\mathcal{U}([0.5, 0.7])$ ,  $\mathcal{U}([0.1, 0.3])$ ,  $\mathcal{U}([0.1, 0.3])$ ,  $\mathcal{U}([0; 0.02])$ . Les résultats sont représentées sur la figure 7.5.

Estimation des paramètres

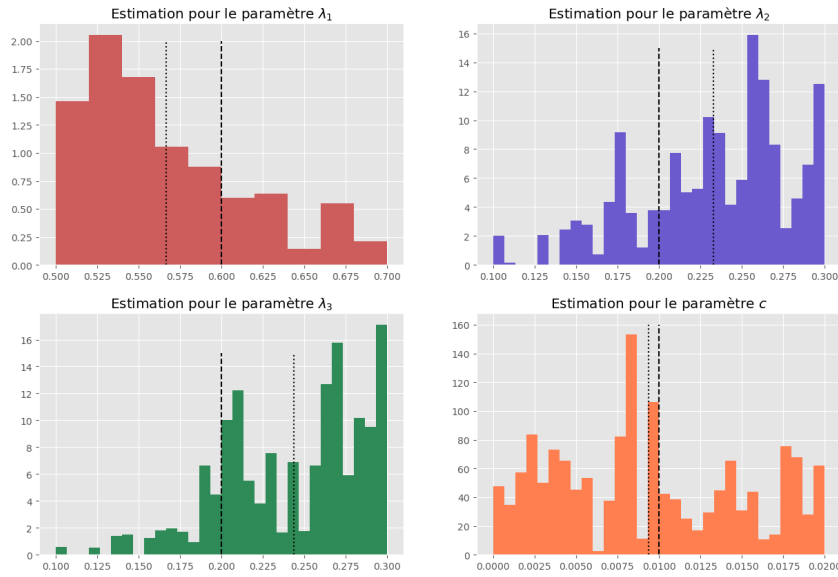


FIGURE 7.5 – Histogramme des différentes valeur prior simulées, pondérées par  $W_i$ , avec la moyenne pondérée des tirages (en pointillet), et les paramètres des données observées (en tiret).

L'estimation du  $c$ , qui est ici notre but, semble être correcte si on considère la valeur moyenne pondérée de la posterior obtenue. Pour  $c_{obs} = 0.01$ , on obtient  $\hat{c}_{moy} = 0.0094$ , avec donc une erreur de  $\varepsilon = 6 \times 10^{-4}$ .

Si nous simulons le processus pour les paramètres moyens obtenus, c'est-à-dire la moyenne de la posterior obtenue, nous obtenons la courbe représenté en figure 7.6. La notations  $\hat{x}_{moy}$  indique la quantité estimée par la moyenne de la posterior. Cette simulation ne semble pas approcher correctement les

données obtenues, la fonction  $S^N$  est sur-estimer, tandis que les autres sont sous-estimés. Cela peut provenir du nombre faible de simulations gardés, avec seulement 400 des simulations qui sont gardés pour estimer la procédure; cela peut également convenir d'un choix trop large pour les lois prior.

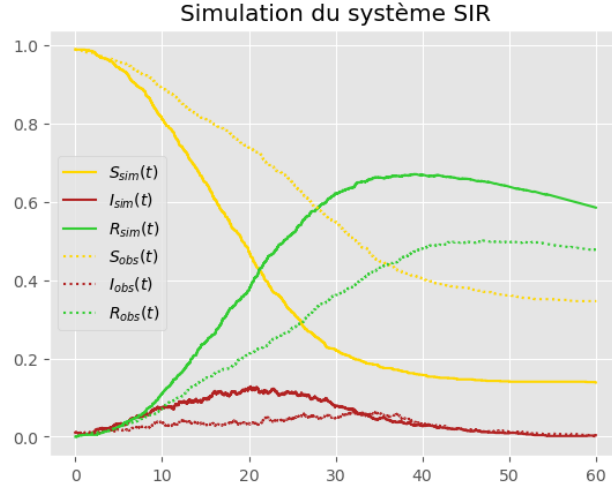


FIGURE 7.6 – Simulation du processus pour  $\hat{\lambda}_{1\text{moy}}, \hat{\lambda}_{2\text{moy}}, \hat{\lambda}_{3\text{moy}}, \hat{c}_{\text{moy}}$

Nous simulons également le processus pour les paramètres maximisant le poids. Les simulations maximisant la valeur du noyau  $K_\delta$  sont celles minimisant la distance à la courbe des données observées. Il est donc possible d'estimer nos paramètres par la valeur des paramètres  $\hat{x}_{\text{max}}$  maximisant la valeur du noyau après simulation. Les résultats de cette simulation sont représentés en figure 7.7. Cette estimation semble mieux que la précédente, mais sur-estime encore le niveau de susceptible, et sous-estime les deux autres quantités.

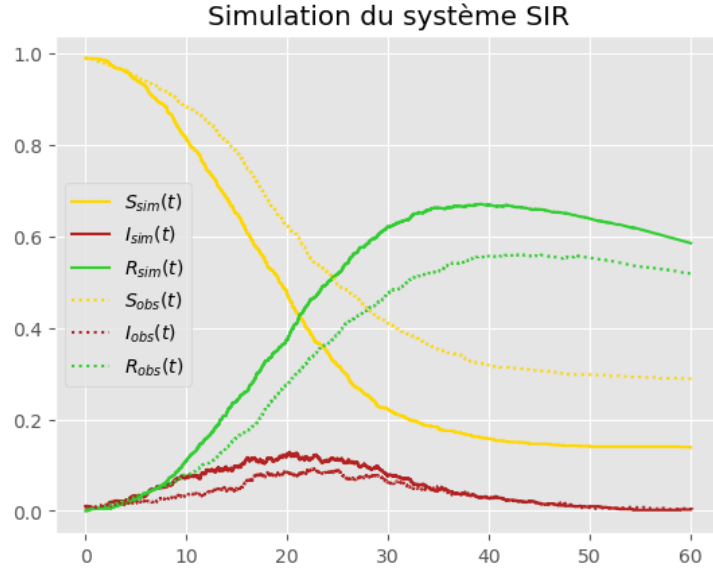


FIGURE 7.7 – Simulation du processus pour  $\hat{\lambda}_{1\text{max}}, \hat{\lambda}_{2\text{max}}, \hat{\lambda}_{3\text{max}}, \hat{c}_{\text{max}}$

---

## Conclusion

---

Nous avons donc implémenté les simulations correspondant au modèle décrit, et l'algorithme ABC pour l'estimation des paramètres. Tout les codes utilisés sont disponibles sur [https://github.com/aurianegbt/HIV\\_stats](https://github.com/aurianegbt/HIV_stats). Au vu des résultats, il semblerait que nous n'ayons pas considéré de bonnes priors, et éventuellement pas fait assez de simulations. Pour corriger le second problème, il est possible d'augmenter les seuils  $\delta_1, \delta_2$  d'acceptation, au risque de garder des simulations moins pertinentes pour l'estimation. Une meilleure solution serait d'augmenter le nombre  $N$  de simulations utilisées, ce qui n'a pas été possible avec le matériel à disposition. Néanmoins, la valeur  $\hat{c}_{moy}$  obtenue comme estimateur de l'efficacité du contact tracing est proche de la valeur utilisée pour les données, et peut donc être considéré comme un estimateur fiable.

---

## Liste des références

---

- Michael G. B. Blum and Viet Chi Tran. HIV with contact tracing : a case study in approximate Bayesian computation. *Biostatistics*, 11(4) :644–660, 05 2010.
- Iacus Stefano M. *Simulation and inference for stochastic differential equations : with R examples*. Berlin ; Heidelberg DEU ; New York ; Bâle CHE : Springer , 2008, 2008.
- S. Méléard. *Modèles aléatoires en Ecologie et Evolution*. Mathématiques et Applications. Springer Berlin Heidelberg, 2016.

# Procédure post-hoc



Auriane GABAUT et Alexandre PERRIN  
Encadré par Guillermo DURAND et Marie CHION  
Novembre 2022 - Mars 2023

## Sommaire

---

|                                                                                                       |            |
|-------------------------------------------------------------------------------------------------------|------------|
| <b>Introduction et Motivations</b> . . . . .                                                          | <b>146</b> |
| <b>8.1 Analyse de données de protéomiques</b> . . . . .                                               | <b>147</b> |
| 8.1.1 L'étude des protéomes . . . . .                                                                 | 147        |
| 8.1.2 Analyse Protéomique par Spectrométrie de masse . . . . .                                        | 148        |
| 8.1.3 Les outils statistiques usuels en protéomique . . . . .                                         | 149        |
| <b>8.2 Procédure Post-Hoc</b> . . . . .                                                               | <b>150</b> |
| 8.2.1 Préliminaires . . . . .                                                                         | 150        |
| 8.2.2 Approche de Blanchard et al. . . . .                                                            | 157        |
| 8.2.3 Approche de Durand et al. . . . .                                                               | 159        |
| 8.2.4 Hybridation de bornes . . . . .                                                                 | 161        |
| <b>8.3 Études de cas : applications des bornes Post-Hoc sur des données de protéomiques</b> . . . . . | <b>162</b> |
| 8.3.1 Un premier jeu de données . . . . .                                                             | 162        |
| 8.3.2 Étude de cas sur un jeu de données contrôlées . . . . .                                         | 167        |
| <b>Conclusion</b> . . . . .                                                                           | <b>176</b> |

---

---

## Introduction et Motivations

---

L'analyse des données de protéomique est une discipline qui vise à identifier et quantifier les différentes protéines présentes dans un échantillon. Dans ce contexte, les expériences génèrent souvent de grandes quantités de données qui nécessitent des outils statistiques pour leur interprétation. Les bornes post hoc sont des outils statistiques qui permettent de garantir la fiabilité des analyses, même dans des cas où les données ont été pré-sélectionnées.

Dans ce projet, notre objectif est d'étudier les bornes post hoc développées par Blanchard, Durand et al. [3, 6] et de les appliquer à des données de protéomique réelles. Pour ce faire, nous introduirons les

notions et enjeux des tests multiples en statistiques. Nous évoquerons ensuite comment les bornes post hoc fournissent un outil statistique fiable et capable de résister à la sélection de données. Ces bornes ont déjà été testées sur des simulations de données, mais notre contribution consiste à les appliquer à des données réelles.

---

## Remerciements

---

Nous tenons à remercier Guillermo et Marie pour leur accompagnement sans faille tout au long du projet. Leur réactivité et leur disponibilité ont été précieuses pour nous permettre de réaliser un travail dont nous sommes pleinement satisfaits.

Nous remercions également Pierre NEUVIAL pour nous avoir consacré du temps afin de discuter de notre projet et de nos différentes approches. Sa contribution nous a permis de mieux cerner les enjeux de notre travail et de l'enrichir.

---

## 8.1. Analyse de données de protéomiques

---

### 8.1.1. L'étude des protéomes

Un protéome se définit comme l'ensemble des protéines codées par un génome. Les protéines sont issues de la traduction des ARN messagers, eux même issus de la transcription de l'ADN. Elles sont des hétéropolymères orientés et séquencés d'acides aminés, capables d'adopter des structures tridimensionnelles spécialisées dans certaines fonctions, essentielles dans le vivant. Elles assurent la structure des cellules, la catalyse de voies métaboliques, la communication intra et extra-cellulaire, l'expression génétique et plus généralement la maintenance des cellules. Constituant un bloc fondamental dans les mécanismes biologiques, la compréhension des rôles de chaque protéine et de leurs interactions est nécessaire pour expliquer les phénomènes biologiques. La protéomique répond en partie à ces questions, et représente une branche de la biologie moléculaire en expansion depuis la fin des années 90.

Une des approches primordiales lors d'une étude biologique consiste en la comparaison de conditions, autrement dit, la règle du "témoin". En effet, une étude n'est conditionnée que par l'obtention d'individus

différenciables par des caractères. Par exemple, on distinguera un individu sain et un individu malade, ou un individu sauvage et un individu muté. Puis, la comparaison entre ces conditions est permise par diverses approches expérimentales. La génomique permet de révéler les mutations génétiques, la transcriptomique permet de révéler les différences d'intensité de transcription et la protéomique permet de révéler les différences de niveau de traduction d'une protéine. A chaque approche est associé une hypothèse propre. L'hypothèse sous-jacente motivant le recours à la protéomique suppose que les différences de concentrations protéiques au sein de deux systèmes biologiques expliquent les différences phénotypiques. Par exemple, on supposera qu'une chenille présente des proportions protéiques différentes de celles de son congénère métamorphosé. La force de cette hypothèse est qu'elle est vraie pour la très grande majorité des cas où l'on observe une différence phénotypique. En effet, le phénotype étant d'origine génétique et environnementale, il s'exprime dans la plupart des cas au travers d'une modulation du protéome. Ainsi, la protéomique apporte une information complémentaire à la génomique car le génome ne prédit pas l'impact environnemental sur le phénotype.

En somme, la protéomique permet d'avoir une vue d'ensemble sur les différentes protéines constituant un système, en identifiant les protéines d'abondances significativement différentes entre plusieurs individus ou échantillons. Cela offre l'avantage de fournir des informations fines sur l'origine des différences phénotypiques, essentielles dans la compréhension des mécanismes biologiques.

### 8.1.2. Analyse Protéomique par Spectrométrie de masse

La spectrométrie de masse est une technique de chimie analytique de choix pour les analyses de protéomique quantitative. Cette méthode requiert une étape préliminaire de purification des échantillons biologiques étudiés. Après purification, les protéines passent par une étape de réduction et d'alkylation, puis sont digérées par des protéases. Cette dernière étape permet de fragmenter les protéines en peptides. Les protéases, étant regiospécifiques, permettent d'obtenir toujours les mêmes fragments avec la même séquence peptidique à partir d'une même protéine. Une fois les peptides purifiés, on réalise une étude par spectrométrie de masse. La première étape consiste à réaliser une première ségrégation des peptides par chromatographie liquide (LC) suivie d'une lecture de leur rapport masse sur charge dans un spectromètre de masse (MS1). Cette étape fournit un spectre répertoriant les différents peptides avec leur intensité. L'expérimentateur peut alors choisir les peptides les plus intenses pour l'étape suivante. La deuxième étape consiste en une fragmentation des peptides sélectionnés suivie d'une lecture de leur spectre de fragmentation par le même spectromètre (MS-MS). Ce spectre permet l'identification de la séquence de chaque peptide en acides aminés. A partir de la séquence peptidique il est possible de déterminer le rapport masse/charge théorique de chaque peptide et de les identifier sur le spectre obtenu lors du MS1. A l'aide de moteurs de recherche spécialisés, la séquence d'un peptide permettra d'identifier ultérieurement la protéine dont il provient. La quantification de chaque peptide est réalisée selon deux méthodes. Une première méthode consiste en une mesure de l'aire sous le pic chromatographique associé au peptide dans le spectre MS1. Une seconde méthode consiste en un comptage de spectres MS-MS, ayant permis l'identification du peptide. Les peptides sont alors quantifiés en intensité proportionnelle à la concentration de leur protéine associée dans les échantillons (figure 8.1).



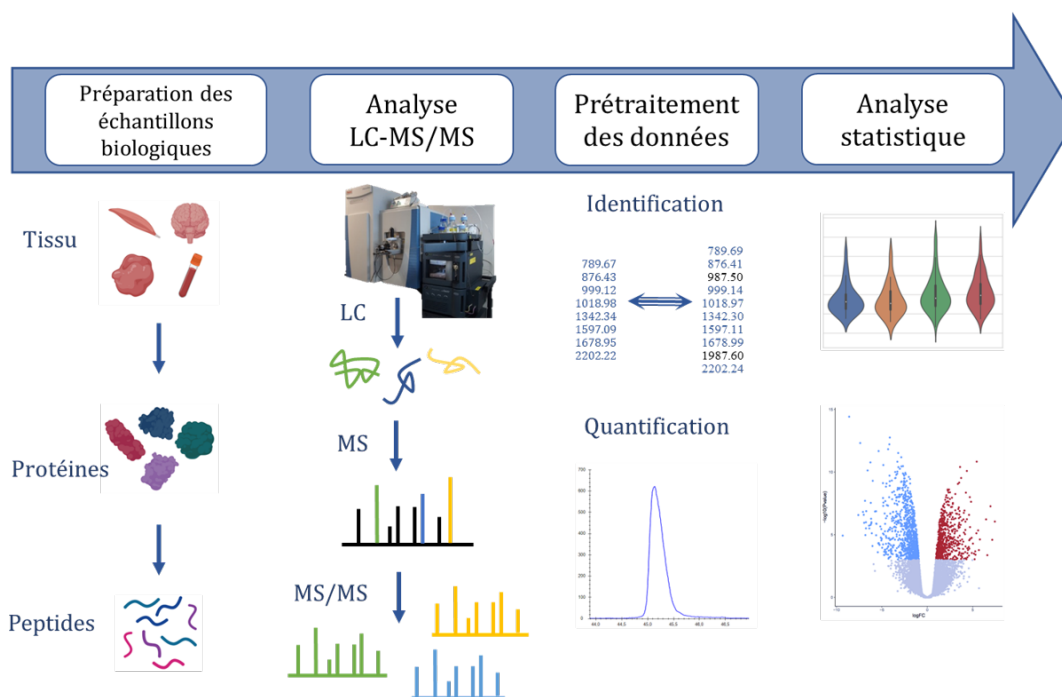


FIGURE 8.1 – Protocole d’une analyse protéomique par spectrométrie de masse

### 8.1.3. Les outils statistiques usuels en protéomique

Les intensités obtenues par la protéomique quantitative sont comparées entre différentes conditions par différents tests statistiques (Test de Student usuellement, détaillé plus formellement en partie 8.2.1). De nombreux outils informatiques sont à la disposition de biologistes afin de sélectionner un premier set de peptides, susceptibles d’apporter des informations [13]. Cette sélection prior se fait notamment par le choix du seuil de significativité du test et du Fold change (ratio des intensités entre les conditions). Les progrès technologiques des dernières décennies permettent aujourd’hui d’identifier et de quantifier des plusieurs milliers de peptides. Selon les concentrations protéiques dans les échantillons, le nombre de peptides recensés est de l’ordre de  $10^4$  à  $10^5$ . Ainsi, lors d’analyses différentielles, où autant de tests que de peptides sont réalisés, le nombre de faux positifs tend à exploser. Par conséquent, un simple test univarié sur l’ensemble des peptides ne garantit pas une sélection satisfaisante des tests significatifs. Les procédures post-hoc sont alors des outils statistiques capables de répondre à ce problème, en fournissant des garanties statistiques robustes à la sélection évoquée. Elles fournissent également une procédure de rejet permettant de limiter le nombre de faux positifs dans les données [8]. La sélection prior d’un ensemble de p-valeurs pouvant amener à un biais, une approche inverse peut être considérée pour fournir une garantie sur les faux positifs dans cet ensemble.

---

## 8.2. Procédure Post-Hoc

---

Pour ce projet, nous nous appuyons sur les travaux effectués par G. DURAND, G. BLANCHARD, P. NEUVIAL et E. ROQUAIN. Nous commencerons par développer quelques notions de statistiques, puis énoncerons les résultats théoriques sur les bornes post hoc. Les démonstrations et idées des résultats sont détaillées dans leurs articles [3, 6] et ne sont pas l'objet de notre projet.

### 8.2.1. Préliminaires

Pour commencer, les procédures post hoc sont utilisées dans le cadre des tests multiples en grande dimension, très fréquents en protéomique. Dans ces études, des centaines, voire des milliers de peptides sont analysés simultanément, et des tests statistiques sont utilisés pour identifier les peptides dont l'expression diffère significativement entre deux conditions.

L'utilisation de tests multiples est particulièrement importante en protéomique, car les données générées peuvent conduire à un grand nombre de comparaisons simultanées, ce qui peut entraîner une augmentation du risque de type I, ou "faux positifs". Sans contrôle des erreurs multiples, il est possible de trouver des peptides différentiellement exprimés qui ne sont en réalité pas significatifs. Ceci correspondrait plutôt à un bruit résultant de nos tests multiples.

En utilisant des tests multiples, nous pouvons ajuster le seuil de significativité voulu pour chaque test pour maintenir un risque de type I global acceptable, généralement fixé à 5%. Les corrections de Bonferroni [4], Holm-Bonferroni [11] et Benjamini-Hochberg [1] sont des méthodes couramment utilisées pour contrôler le risque de type I dans l'analyse de données de protéomique.

### Tests Statistiques Simples

Les tests statistiques permettent de valider ou non des hypothèses et de donner un cadre à notre décision, offrant ainsi des garanties statistiques sur la décision prise à partir des données observées  $\mathbf{X} = (X_1, \dots, X_n)$ . Formellement, un test statistique est un test de deux hypothèses, notées  $H_0$  et  $H_1$ . L'hypothèse  $H_0$  est appelée hypothèse nulle, tandis que  $H_1$  est l'hypothèse alternative. Dans le cadre paramétrique, c'est-à-dire avec un modèle paramétré par  $\theta \in \Theta \subset \mathbb{R}^d$ , le test peut être réécrit comme suit :

$$H_0 : \theta \in \Theta_0 \text{ vs. } H_1 : \theta \in \Theta_1$$

où  $\Theta_0, \Theta_1 \subset \Theta$  sont disjoints. Ce test est noté  $\phi(\mathbf{X}) \in \{0, 1\}$ , avec  $\phi(\mathbf{X}) = 0$  lorsque  $H_0$  est acceptée, et 1 sinon.

En pratique, la réponse à la question posée par le test est inconnue, ce qui crée une incertitude quant à la décision prise. C'est pour cette raison que le risque associé à chaque décision est également pris en compte. En contrôlant le risque encouru pour prendre une décision, on fournit une première garantie

statistique sur l'étude. Pour discuter de ce niveau de confiance, on introduit les définitions suivantes :

**Définition 8.2.1 -**

Soit l'espace probabilisé  $(\Omega, \mathcal{A}, (\mathbb{P}_\theta)_{\theta \in \Theta})$ , et une variable aléatoire  $X$  à valeurs dans  $\mathbb{R}^d$  de loi supposée  $\mathcal{P} \in (\mathbb{P}_\theta)_{\theta \in \Theta}$ . Soit un échantillon  $\mathbf{X} = (X_1, \dots, X_n) \in \mathbb{R}^d \sim \mathcal{P}$ , observations de  $X$ . Soit le test :

$$H_0 : \theta \in \Theta_0 \quad \text{vs.} \quad H_1 : \theta \in \Theta_1$$

- Lorsque  $H_0$  est rejetée, alors qu'en réalité elle était vraie, on parle d'une erreur de type I, ou de première espèce.
- Lorsque  $H_0$  est conservée, alors qu'en réalité elle était fausse, on parle d'une erreur de type II, ou de seconde espèce.
- La taille du test est définie comme  $\sup_{\theta \in \Theta_0} \mathbb{E}_\theta[\phi]$  et le test est de niveau  $\alpha \in [0, 1]$  si sa taille est inférieure ou égale à  $\alpha$
- La fonction puissance du test est la fonction  $\pi : \theta \in \Theta_1 \mapsto \mathbb{E}_\theta[\phi]$ .

La notation  $E_\theta[\cdot]$  correspond à l'espérance prise dans le cas où  $X \sim \mathbb{P}_\theta$ .

La notion de niveau  $\alpha$  permet de contrôler le risque de type I du test, c'est-à-dire la probabilité de rejeter à tort l'hypothèse nulle  $H_0$ , alors qu'elle est vraie. La puissance du test, quant à elle, contrôle plutôt le risque de type II, c'est-à-dire la probabilité de ne pas rejeter l'hypothèse nulle alors qu'elle est fausse. L'enjeu derrière les statistiques est de trouver un compromis entre ces deux types d'erreurs, en privilégiant l'un ou l'autre en fonction du contexte.

La p-valeur est une variable aléatoire  $p(\mathbf{X})$   $\sigma(\mathbf{X})$ -mesurable. Elle mesure l'évidence contre l'hypothèse nulle  $H_0$  en étant la probabilité d'obtenir un résultat au moins aussi peu vraisemblable que celui observé, si  $H_0$  est vraie. De manière générale, plus la p-valeur est faible, plus les résultats suggèrent de rejeter l'hypothèse nulle. En pratique, si la p-valeur est très faible, en particulier inférieure au seuil de significativité fixé, les données observées sont très peu probables sous l'hypothèse nulle, et nous pouvons donc décider de rejeter  $H_0$  au niveau de confiance  $\alpha$ . La p-valeur vérifie la propriété suivante :

**Proposition 8.2.0 -**

$p(\mathbf{X})$  est une variable aléatoire  $\sigma(\mathbf{X})$ -mesurable à valeurs dans l'intervalle  $[0 ; 1]$ , telle que

$$\sup_{\theta \in \Theta_0} \mathbb{P}_\theta(p(\mathbf{X}) \leq u) \leq u, \quad \forall u \in [0 ; 1].$$

Autrement dit,  $\mathbb{P}_\theta(p(\mathbf{X}) \leq \alpha) \leq \alpha$  pour tout  $\theta \in \Theta_0$ . Cette proposition précise que, pour toute valeur  $u$  dans l'intervalle  $[0 ; 1]$ , la probabilité que la réalisation de la p-valeur  $p(\mathbf{X})$  soit inférieure ou égale à  $u$ , sous l'hypothèse nulle  $H_0$ , est elle-même inférieure ou égale à  $u$ . Finalement, la propriété 1 implique que si la p-valeur  $p(\mathbf{X})$  est inférieure à un niveau de significativité donné  $\alpha$ , alors on peut rejeter l'hypothèse nulle  $H_0$  au niveau de confiance  $\alpha$ . Le rejet de l'hypothèse  $H_0$  est donc équivalent à  $p(\mathbf{X}) \leq \alpha$ .

**Test de Student**

Le test de Student [9], ou t-test, est l'un des tests les plus couramment utilisés en protéomique. Il s'agit d'un test paramétrique qui peut être écrit comme  $\phi(\mathbf{X}) = \mathbb{1}_{|T(\mathbf{X})| > s}$ , où  $s$  est un seuil ajusté en fonction du niveau de confiance souhaité, et  $T(\mathbf{X})$  est la statistique de test. La distribution de cette statistique sous l'hypothèse nulle  $H_0$  suit une loi de Student.

**Définition 8.2.2 -**

Soit  $(X_i)_{1 \leq i \leq k+1}$  une suite de variables aléatoires indépendantes et identiquement distribuées selon une loi normale centrée réduite  $\mathcal{N}(0, 1)$ . La variable aléatoire

$$U = \sum_{i=1}^k X_i^2$$

suit une loi du  $\chi^2$  à  $k$  degrés de liberté, notée  $\chi^2(k)$  et celle

$$T = \frac{X_{k+1}}{\sqrt{\frac{U}{k}}}$$

suit une loi de Student à  $k$  degrés de liberté, notée  $\mathcal{T}(k)$  ou  $t(k)$ .

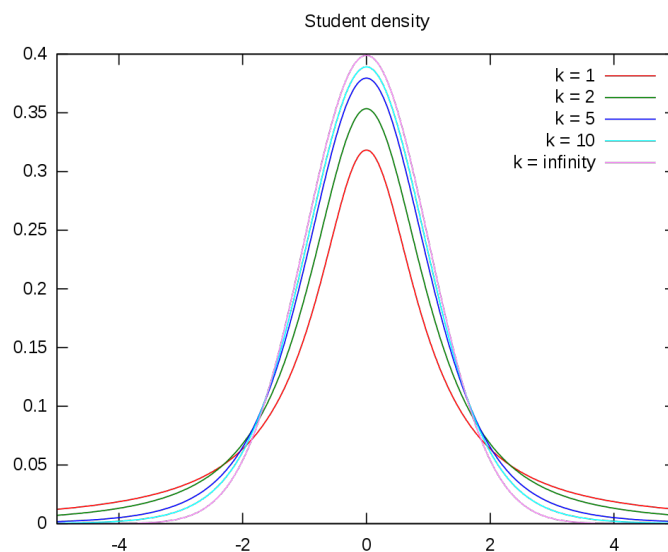


FIGURE 8.2 – Densité des lois de Student pour différentes valeurs de  $k$ .

Lorsque le paramètre  $k$  de la loi de Student tend vers l'infini, la distribution converge vers celle d'une loi normale centrée réduite. Cette propriété est souvent utilisée dans les tests d'hypothèse pour les petits échantillons, où la loi de Student est utilisée pour estimer la distribution de la moyenne d'une population normalement distribuée lorsque l'écart-type est inconnu. Lorsque la taille de l'échantillon est suffisamment grande, la distribution de la moyenne suit approximativement une distribution normale, ce qui permet d'utiliser la loi normale pour les tests d'hypothèse.

En biologie, la loi normale est souvent utilisée comme modèle théorique pour décrire le comportement de phénomènes biologiques. Il est donc fréquent de considérer un modèle avec  $\Omega = \mathbb{R}$  et  $(P_\theta)_{\theta \in \mathbb{R}} = (\mathcal{N}(\mu, \sigma^2))_{\mu \in \mathbb{R}, \sigma \in \mathbb{R}_+}$ , où donc nos observations sont des réels  $x_1, \dots, x_n$ , réalisations d'une loi normale de paramètres  $\theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+$  inconnus.

Sous ce modèle, un premier exemple pourrait considérer une variable aléatoire  $X$  dans  $\mathbb{R}$  suivant une loi normale de paramètres  $\theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+$  inconnus, dont on observe un échantillon  $\mathbf{X} = (X_1, \dots, X_n)$ . Ces observations pourraient être la taille de plantes d'une certaine espèce, la durée de vie d'une cellule, etc. Pour ces hypothèses, la statistique

$$T(\mathbf{X}) = \sqrt{n} \frac{\bar{X} - \mu_0}{S(\mathbf{X})}$$

avec  $S$  l'estimateur empirique sans biais de la variance  $\sigma^2$  et  $\bar{X}$  la moyenne empirique des observations, suit une loi de Student à  $k = n - 1$  degrés de liberté lorsque  $\mu = \mu_0$ . Finalement, le test  $\phi(\mathbf{X}) \mapsto \mathbb{1}_{|T(\mathbf{X})| > s}$  de

$$H_0 : \mu = \mu_0 \quad \text{vs.} \quad H_1 : \mu \neq \mu_0$$

pour un certain  $\mu_0 \in \mathbb{R}$  est appelé test de Student. On remarque de plus que sous  $H_0$  :

$$\begin{aligned} \mathbb{E}[\phi(\mathbf{X})] &= \mathbb{P}(|T(\mathbf{X})| > s) \\ &= 2\mathbb{P}(T(\mathbf{X}) > s) && \text{par symétrie} \\ &= 2(1 - F_{t_k}(s)) \end{aligned}$$

avec  $F_{t_k}$  la fonction de répartition d'une loi de Student à  $k$  degrés de liberté. Le test  $\phi$  est de niveau  $\alpha \in [0 ; 1]$  lorsque  $s$  est le quantile d'ordre  $1 - \alpha/2$  de  $\mathcal{T}(k)$ , noté  $t_{k,1-\alpha/2}$ . Ainsi, si les données observées donnent  $|T(\mathbf{X})|$  supérieure à  $t_{k,1-\alpha/2}$ , alors on rejette l'hypothèse nulle  $H_0$  et on conclut que la moyenne  $\mu$  est significativement différente de  $\mu_0$  au niveau de confiance  $\alpha$ .

Un deuxième exemple de modélisation pourrait considérer deux conditions  $Cond_A$  et  $Cond_B$ , et des observations  $\mathbf{X}^A = (X_1^A, \dots, X_n^A)$  (resp.  $\mathbf{X}^B = (X_1^B, \dots, X_m^B)$ ) de la condition  $Cond_A$  (resp.  $Cond_B$ ), supposées indépendantes de loi  $\mathcal{N}(\mu_A, \sigma_A^2)$  (resp.  $\mathcal{N}(\mu_B, \sigma_B^2)$ ). Les paramètres inconnus seraient ici  $\theta = (\mu_A, \mu_B, \sigma_A^2, \sigma_B^2)$ . La statistique de test

$$T(\mathbf{X}) = \frac{\bar{X}^A - \bar{X}^B}{\sqrt{\frac{S^A(\mathbf{X})^2}{n} + \frac{S^B(\mathbf{X})^2}{m}}}$$

suit, sous l'hypothèse nulle  $\mu_A = \mu_B$ , une loi de Behrens-Fisher, qui approche alors une loi de Student [12]. Ainsi, le test  $\phi(\mathbf{X}) = \mathbb{1}_{|T(\mathbf{X})| > s}$  pour tester l'hypothèse  $H_0 : \mu_A = \mu_B$  contre l'hypothèse alternative  $H_1 : \mu_A \neq \mu_B$  est un test de Student, également appelé t-test de Welch [17]. Dans le cas particulier où les deux conditions sont supposées de même variance, la statistique de test s'écrit

$$T(X) = \frac{\bar{X}^A - \bar{X}^B}{\sqrt{S(\mathbf{X})^2 \left(\frac{1}{n} + \frac{1}{m}\right)}}$$

et suit une loi de Student sous l'hypothèse  $H_0$ .

En résumé, le test de Student classique assume que les variances des deux populations sont égales, c'est-à-dire  $\sigma_A^2 = \sigma_B^2$ . Si cette hypothèse n'est pas vérifiée, on peut utiliser le t-test de Welch, qui est basé sur une approximation de la loi de la statistique de test par une loi de Student généralisée. Cette approche est plus robuste et peut être utilisée même si les variances des deux populations sont différentes.

### Les enjeux des tests multiples

Nous souhaitons effectuer simultanément  $m > 1$  tests. Chaque test  $i \in \{1, \dots, m\}$  s'écrit comme un test de  $H_{0,i}$  contre  $H_{1,i}$ . À partir de données  $\mathbf{X} = (X_1, \dots, X_n)$ , nous obtenons une p-valeur pour chacun de ces tests, notée  $p_i(\mathbf{X})$ . Dans la suite, nous utiliserons la notation  $p_i$  pour simplifier. Nous définissons également l'ensemble  $\mathcal{H}_0 \subset \{1, \dots, m\}$  des hypothèses  $H_{0,i}$  qui sont vraies, c'est-à-dire  $\mathcal{H}_0 = \{i \in \{1, \dots, m\}, H_{0,i} \text{ est vraie}\}$ ; et la fonction

$$V : S \subset \{1, \dots, m\} \mapsto |S \cap \mathcal{H}_0|.$$

En pratique, l'ensemble  $\mathcal{H}_0$  n'est pas connu.

Le cadre des tests simples ne peut pas s'appliquer ici. Si nous imposons un niveau  $\alpha$  pour chaque test  $i \in \{1, \dots, m\}$ , comme nous le ferions habituellement, nous surestimons alors en moyenne le nombre d'hypothèses significatives, c'est-à-dire d'hypothèses fausses. En effet, d'après la proposition 1 pour tout  $i \in \{1, \dots, m\}$ ,  $\mathbb{P}_{H_{0,i}}(p_i \leq \alpha) \leq \alpha$  pour tout  $\alpha \in [0; 1]$ , et nous obtenons alors, pour une région de rejet  $R$  de la forme  $R = \{i \in \{1, \dots, m\} : p_i \leq \alpha\}$  :

$$\begin{aligned} \mathbb{E}[V(R)] &= \sum_{i \in \mathcal{H}_0} \mathbb{P}_{H_{0,i}}(p_i \leq \alpha) \\ &\leq \sum_{i \in \mathcal{H}_0} \alpha \\ &\leq \alpha m_0 \end{aligned}$$

où  $m_0 = |\mathcal{H}_0|$ . Ainsi, pour un grand nombre d'hypothèses vraies  $m_0$  testées, la probabilité de faire une erreur en rejetant une hypothèse, c'est-à-dire d'avoir un faux positif, peut être grande. Il semble donc essentiel de contrôler les effets de la multiplicité.

Il existe différentes approches possibles pour cela. Le contrôle du nombre de faux positifs peut être réalisé en contrôlant le Family Wise Error Rate (FWER) ou le taux de faux positifs (*False Discovery Rate*, FDR), notamment.

#### Définition 8.2.3 -

Soit un ensemble  $S \subset \{1, \dots, m\}$  d'hypothèses. Le FWER (Family Wise Error Rate) est défini par :

$$FWER(S) = \mathbb{P}(V(S) > 0) \quad (8.1)$$

et le FDR (False Discovery Rate) par :

$$FDR(S) = \mathbb{E} \left[ \frac{V(S)}{|S| \vee 1} \right] \quad (8.2)$$

Nous définissons également le FDP (False Discovery Proportion) par :

$$FDP(S) = \frac{V(S)}{|S| \vee 1}$$

de sorte que  $FDR(S) = \mathbb{E}[FDP(S)]$ .

#### Proposition 8.2.0 -

Nous avons toujours  $FDR(S) \leq FWER(S)$ .

Les deux approches, FWER et FDR, peuvent être plus ou moins pertinentes selon le cadre de l'étude. Le FWER permet de contrôler la présence de faux positifs dans la région de rejet choisie, tandis que le FDR contrôle la proportion de faux positifs parmi les hypothèses rejetées. Ainsi, le choix entre FWER et FDR peut dépendre des objectifs de l'étude et du niveau de risque acceptable de faux positifs. L'idée derrière le contrôle du FWER est que l'on ne veut absolument pas de faux positifs, ce qui en fait une méthode préférable dans les études où il est crucial de minimiser le risque de faux positifs dans l'ensemble des hypothèses testées. Le FDR, en revanche, est plus approprié pour les études où il est acceptable d'avoir quelques faux positifs, tant que leur nombre reste contrôlé.

### Procédures de tests multiples

Nous allons maintenant étudier quelques procédures permettant de contrôler le FDR (8.2) ou le FWER (8.1). La première procédure évoquée est celle de Bonferroni [4], suivie de la procédure de Holm-

Bonferroni [11] et enfin celle de Benjamini-Hochberg [1].

**Procédure de Bonferroni.** La correction de Bonferroni définit la région de rejet par :

$$R_B = \left\{ i \in \{1, \dots, m\} : p_i \leq \frac{\alpha}{m} \right\}$$

Pour cette région, le FWER est borné par  $\alpha$  :

$$\begin{aligned} FWER(R_B) &= \mathbb{P} \left( \exists i \in \mathcal{H}_0 : p_i \leq \frac{\alpha}{m} \right) \\ &\leq \sum_{i \in \mathcal{H}_0} \mathbb{P}_{H_{0,i}} \left( p_i \leq \frac{\alpha}{m} \right) \\ &\leq \alpha \frac{m_0}{m} \leq \alpha \end{aligned}$$

où  $m_0 = |\mathcal{H}_0|$ .

Cette méthode est très conservative, mais elle peut être trop restrictive dans les études où le nombre d'hypothèses testées est élevé.

**Procédure de Holm-Bonferroni.** Soit  $(p_{(k)})_{1 \leq k \leq m}$  la suite ordonnée par ordre croissant des p-valeurs  $(p_i)_{1 \leq i \leq m}$ . Nous posons  $\hat{k} = \min \left\{ k \in \{1, \dots, m\} : p_{(k)} > \frac{\alpha}{m - k + 1} \right\}$ . Alors, la procédure de Holm-Bonferroni est définie par :

$$R_{HB} = \left\{ i \in \{1, \dots, m\} : k \leq \hat{k} \text{ et } p_i \leq \frac{\alpha}{m - k + 1} \right\}$$

où  $k$  est le rang dans  $(p_{(k)})_k$  de  $p_i$ . Cette correction contrôle le FWER au niveau  $\alpha$ .

Dans le cas où l'on rejette au moins une hypothèse vraie, si

$$h = \min \{ i \in \{1, \dots, m\} : H_{0,(i)} \text{ rejetée mais vraie} \}$$

alors les hypothèses  $H_{0,(1)}, \dots, H_{0,(h-1)}$  sont correctement rejetées. D'où :

$$\begin{aligned} \mathcal{H}_0 &\subset \{1, \dots, m\} \setminus \{(1), \dots, (h-1)\} \\ |\mathcal{H}_0| &\leq m - h + 1 \\ m_0 &\leq m - h - 1 \end{aligned}$$

De plus, nous avons  $p_{(h)} \leq \frac{\alpha}{m - h + 1} \leq \frac{\alpha}{m_0}$ , et donc, dans l'hypothèse que l'on rejette au moins une hypothèse vraie, il existe une hypothèse  $i \in \mathcal{H}_0$  telle que  $p_i \leq \frac{\alpha}{m_0}$ . Nous avons finalement :

$$\begin{aligned} FWER(R_{HB}) &= \mathbb{P}(V(R_{HB}) > 1) = \mathbb{P}(\exists i \in R_{HB}, i \in \mathcal{H}_0) \\ &\leq \mathbb{P} \left( \exists i \in \mathcal{H}_0, p_i \leq \frac{\alpha}{m_0} \right) \\ &\leq \sum_{i \in \mathcal{H}_0} \mathbb{P}_{H_{0,i}} \left( p_i \leq \frac{\alpha}{m_0} \right) \\ &\leq \alpha \end{aligned}$$

La correction de Holm-Bonferroni est moins conservative que la correction de Bonferroni.

**Procédure de Benjamini-Hochberg.** La procédure de Benjamini-Hochberg repose sur une hypothèse de "Dépendance de Régression Positive Faible" ou *Weak Positive Regression Dependency Property* (WPRD) en anglais.

**Définitions 8.2.4 -**

Les p-valeurs  $(p_i)_{1 \leq i \leq m}$  respectent l'hypothèse WPRD si pour toute fonction  $g : [0 ; 1]^m \rightarrow \mathbb{R}_+$  mesurable et croissante, nous avons pour  $i \in \mathcal{H}_0$  :

$$\forall s, t \in \{u \in [0 ; 1] : \mathbb{P}(p_i \leq u) > 0\}, s \leq t \Rightarrow \mathbb{E}[g(p_i)|p_i \leq s] \leq \mathbb{E}[g(p_1, \dots, p_m)|p_i \leq t]$$

Autrement dit, la fonction

$$u \in \{u \in [0 ; 1] : \mathbb{P}(p_i \leq u) > 0\} \mapsto \mathbb{E}[g(p_i)|p_i \leq u]$$

est croissante pour tout  $i \in \mathcal{H}_0$ .

*Remarque.* Une famille de p-valeurs indépendantes vérifie la propriété WPRD.

Dans l'hypothèse WPRD, la procédure de Benjamini-Hochberg définit :

$$R_{BH} = \left\{ i \in \{1, \dots, m\} : p_i \leq \frac{\alpha \hat{k}}{m} \right\}, \text{ avec } \hat{k} = \max \left\{ k \in \{1, \dots, m\} : p_{(k)} \leq \frac{\alpha k}{m} \right\}$$

avec la convention que  $R_{BH} = \emptyset$  si  $\{k \in \{1, \dots, m\} : p_{(k)} \leq \alpha k/m\} = \emptyset$ . Cette procédure contrôle alors le FDR au niveau  $\alpha$ .

En effet, nous remarquons que  $\hat{k} = |R_{BH}|$ , d'où :

$$\begin{aligned} FDR(R_{BH}) &= \mathbb{E} \left[ \frac{V(R_{BH})}{|R_{BH}| \vee 1} \right] \\ &= \mathbb{E} \left[ \frac{V(R_{BH})}{\hat{k} \vee 1} \right] \\ &= \sum_{i \in \mathcal{H}_0} \mathbb{E} \left[ \frac{\mathbb{1}_{\hat{k} \geq 1}}{\hat{k}} \mathbb{1}_{p_i \leq \alpha \hat{k}/m} \right] \\ &= \sum_{i \in \mathcal{H}_0} \sum_{k=1}^m \frac{1}{k} \mathbb{P} \left( \{\hat{k} = k\} \cap \left\{ p_i \leq \frac{\alpha k}{m} \right\} \right) \end{aligned}$$

Or  $\mathbb{P} \left( p_i \leq \frac{\alpha k}{m} \right) \leq \frac{\alpha k}{m}$ , donc si nous posons la convention  $\mathbb{P} \left( \hat{k} = k \mid p_i \leq \frac{\alpha k}{m} \right) \mathbb{P} \left( p_i \leq \frac{\alpha k}{m} \right) = 0$  lorsque  $\mathbb{P} \left( p_i \leq \frac{\alpha k}{m} \right) = 0$ , nous obtenons :

$$\begin{aligned} FDR(R_{BH}) &\leq \sum_{i \in \mathcal{H}_0} \sum_{k=1}^m \frac{\alpha}{m} \mathbb{P}(\hat{k} = k | p_i \leq \alpha k/m) \\ &\leq \frac{\alpha}{m} \sum_{i \in \mathcal{H}_0} \sum_{k \leq m} \frac{\alpha}{m} \left( \mathbb{P}(\hat{k} \leq k | p_i \leq \alpha k/m) - \mathbb{P}(\hat{k} \leq k-1 | p_i \leq \alpha k/m) \right) \\ (WPRD) &\leq \frac{\alpha}{m} \sum_{i \in \mathcal{H}_0} \sum_{k \leq m} \frac{\alpha}{m} \mathbb{P} \left( \hat{k} \leq k \mid p_i \leq \frac{\alpha(k+1)}{m} \right) - \mathbb{P} \left( \hat{k} \leq (k-1) \mid p_i \leq \frac{\alpha((k-1)+1)}{m} \right) \\ (téléscopage) &\leq \frac{\alpha}{m} \sum_{i \in \mathcal{H}_0} \mathbb{P} \left( \hat{k} \leq m \mid p_i \leq \frac{\alpha(m+1)}{m} \right) \\ &\leq \alpha \end{aligned}$$

Cette procédure est plus puissante que les deux précédentes mais exige une propriété de dépendance faible.



### 8.2.2. Approche de Blanchard et al.

Nous avons présenté quelques procédures de rejet dans le cadre des tests multiples. Ces procédures sont très utilisées par les biologistes, mais parfois imprudemment, ce qui peut entraîner la perte des garanties statistiques pourtant fondamentales. En effet, une fois les tests multiples effectués sur un grand nombre  $m$  d'hypothèses, il peut arriver que seulement un sous-ensemble d'hypothèses testées soit ensuite étudié, créant ainsi un biais de sélection. Les procédures post hoc ont pour objectif de construire des garanties statistiques résistantes à cette sélection, quelles que soient les données.

Pour atteindre cet objectif, il est nécessaire de construire des bornes appelées bornes post hoc, qui permettent d'obtenir ces garanties et de contrôler le FDP, le FDR ou le FWER, définis précédemment (8.1), (8.2).

#### Définition 8.2.5 -

Une borne post hoc de niveau  $\alpha$  est une fonction  $\hat{V}$  définie sur n'importe quel sous-ensemble  $S$  d'hypothèses de  $\{1, \dots, m\}$  tel que :

$$\mathbb{P}(\forall S \subset \{1, \dots, m\} : |\mathcal{H}_0 \cap S| \leq \hat{V}(S)) \geq 1 - \alpha \quad (8.3)$$

Une borne post hoc fournit donc une borne sur le nombre de faux positifs, et donc le FDP, de tous les ensembles ensemble avec une grande probabilité. Les bornes que nous étudions pour ce projet sont construites à partir de la notion de famille de référence proposée par Blanchard et al. [3].

### Familles de références

Pour construire ces bornes, une méthode est proposée par Blanchard et al. [3], utilisant la notion de famille de références.

#### Définition 8.2.6 - Famille de références

On appelle famille de références une famille

$$\mathfrak{R} = ( \underbrace{\mathcal{R}_k}_{\subset \{1, \dots, m\}}, \underbrace{\zeta_k}_{\in \mathbb{N}} )_{k \in \mathcal{K}}$$

telle que :

$$\text{JER}(\mathfrak{R}) := \mathbb{P}(\exists k \in \mathcal{K}, |\mathcal{R}_k \cap \mathcal{H}_0| > \zeta_k) \leq \alpha$$

où JER est le Joint Error Rate de la famille de références.

Blanchard et al. montrent que, muni d'une telle famille, il est alors possible de construire des bornes post hoc de niveau  $\alpha$  par interpolation lorsque le JER est lui-même contrôlé au seuil  $\alpha$ . Nous allons présenter ces bornes, dans l'optique de tester leur performance sur des données réelles en énonçant leurs propriétés.

#### Définition 8.2.7 -

Soit  $S \subset \{1, \dots, m\}$ , et  $\mathfrak{R} = (R_k, \zeta_k)_{k \leq K}$  une famille de référence. La borne  $V_{\mathfrak{R}}^*$  est défini par

$$V_{\mathfrak{R}}^*(S) = \max_{A \in \mathcal{A}(\mathfrak{R})} |S \cap A| \quad (8.4)$$

avec  $\mathcal{A}(\mathfrak{R}) = \{A \subset \{1, \dots, m\} : \forall k \in \{1, \dots, K\}, |R_k \cap A| \leq \zeta_k\}$ ,

L'ensemble  $\mathcal{H}_0$  des vraies hypothèses est dans  $\mathcal{A}(\mathfrak{R})$ . Cette première borne  $V_{\mathfrak{R}}^*$  est une borne post hoc

et elle est optimale. Elle n'est pourtant pas calculable en général, dès lors que  $m$  est trop grand, puisque le cardinal des sous-ensembles de  $\{1, \dots, m\}$  est de  $2^m$ . Il est donc nécessaire de construire d'autres bornes, qui feraient un compromis entre calculable et optimal.

*Définition 8.2.8 -*

Soit  $S \subset \{1, \dots, m\}$ , et  $\mathfrak{R} = (R_k, \zeta_k)_{k \leq K}$  une famille de référence. On définit la borne post hoc  $\bar{V}_{\mathfrak{R}}$  par

$$\bar{V}_{\mathfrak{R}}(S) = \min_{k \in \{1, \dots, m\}} (|S \setminus R_k| + \zeta_k) \wedge |S| \quad (8.5)$$

La borne  $\bar{V}_{\mathfrak{R}}$  est cette fois calculable, mais on a toujours  $V_{\mathfrak{R}}^* \leq \bar{V}_{\mathfrak{R}}$  du fait de l'optimalité de  $V_{\mathfrak{R}}^*$ . Cependant, dans le cas particulier d'une famille de référence satisfaisant l'hypothèse d'emboîtement

$$\forall k, k' \leq K, R_k \subset R_{k'} \text{ ou } R_{k'} \subset R_k, \quad (8.6)$$

nous avons alors  $\bar{V}_{\mathfrak{R}} = V_{\mathfrak{R}}^*$ .

### Un cas particulier de familles de références

Nous allons voir un cas particulier de familles de référence, définies à partir de région  $(R_k)_{k \leq K}$  de la forme  $R_k = \{i : p_i \leq t_k\}$  où  $\mathcal{T} = (t_k)_{k \leq K}$  est une famille croissante de réels de  $[0 ; 1]$ . Pour ces types de familles, le JER est donné par :

$$JER(\mathcal{T}) = \mathbb{P}(\exists k \in \{1, \dots, K\} : p_{(k; H_0)} < t_k) \quad (8.7)$$

où  $p_{(k, H_0)}$  est la  $k^e$  plus petite p-valeur de  $\{p_i : i \in \mathcal{H}_0\}$ . Nous définissons ensuite une famille de bornes post hoc à partir de ces familles lorsqu'elle contrôle le JER par un réel  $\alpha \in [0 ; 1]$ .

*Définition 8.2.9 -*

Soit  $S \subset \{1, \dots, m\}$  un sous-ensemble d'hypothèses, et une famille  $\mathcal{T} = (t_k)_{k \leq K}$  qui contrôle le JER (8.7) au taux  $\alpha$ . Alors la fonction :

$$V_{\mathcal{T}}(S) = \min_{k \in \{1, \dots, K\}} \left\{ \sum_{i \in S} \mathbb{1}_{\{p_i(X) \geq t_k\}} + k - 1 \right\}$$

est une borne post hoc.

Dans le cas particulier de la famille Simes  $\mathcal{T}_{simes} = (\alpha k/m)_{k \leq m}$ , le JER est contrôlé par  $\alpha \in [0 ; 1]$  sous certaines hypothèses de dépendance (PRDS) [2]. On appelle cette borne la borne Simes, notée  $V_{simes}$  par la suite.

### $\lambda$ -Calibration

Les bornes ainsi définies offrent des garanties statistiques robustes à la sélection des hypothèses. Néanmoins, elles peuvent être sensibles au contrôle du JER (8.7), qui peut ne pas tenir dans certains cas de dépendance dans les données. Une seconde méthode, résistante à cette dépendance, est alors proposée [3], sous le nom de  $\lambda$ -Calibration. Elle consiste à considérer une famille de référence paramétrée par  $\lambda \in [0 ; 1]$  que l'on ajuste pour permettre le contrôle du JER. Blanchard et al. proposent une méthode algorithmique permettant de construire ce  $\lambda$  cible assurant le contrôle du JER. Ainsi, on peut calculer les bornes énoncées précédemment avec le contrôle adéquat de l'erreur.

### 8.2.3. Approche de Durand et al.

La seconde méthode étudiée pour ce projet, construite par Durand et al. [6], repose sur l'hypothèse d'une structure hiérarchique dans les données. Elle suppose pouvoir construire une famille de référence  $\mathfrak{R} = (R_k, \zeta_k)_{k \leq K}$  telle que

$$\forall k, k' \in \{1, \dots, K\}, \quad R_k \subset R_{k'} \quad \text{ou} \quad R_{k'} \subset R_k \quad \text{ou} \quad R_{k'} \cap R_k = \emptyset. \quad (8.8)$$

Cette hiérarchie décrit une structure en arbre, avec des ensembles d'hypothèses, disjoints sur un même niveau, régionnant nos hypothèses selon le contexte. Pour une expérience de protéomique par exemple, un régionnement naturel est de considérer chaque peptide individuellement, ce serait nos feuilles, puis chaque protéine, composée de peptides, qui formerait un deuxième niveau, et enfin, on pourrait éventuellement considérer les gènes regroupant les protéines. Ces régions sont déterminées dans ces structures et nous pouvons les représenter comme sur la figure 8.3.

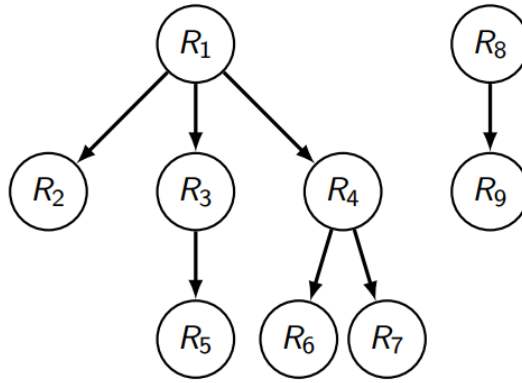


FIGURE 8.3 – Exemple de structure en arbre.

### Structure en forêt

Nous introduisons une borne intermédiaire :

*Définition 8.2.10 -*

Soit  $S \subset \{1, \dots, m\}$  un sous-ensemble d'hypothèses testées,  $\mathfrak{R} = (R_k, \zeta_k)_{k \leq K}$  une famille de référence et  $q \in \mathbb{N}^*$  tel que  $q \leq K$ . Nous définissons la fonction suivante

$$\tilde{V}_{\mathfrak{R}}^q(S) = \min_{\substack{Q \subset \{1, \dots, K\}, \\ |Q| \leq q}} \left( \sum_{k \in Q} \zeta_k \wedge |S \cap R_k| + |S \setminus \cup_{k \in Q} R_k| \right) \quad (8.9)$$

Pour  $q = K$ , nous notons  $\tilde{V}_{\mathfrak{R}} = \tilde{V}_{\mathfrak{R}}^q$ .

*Proposition 8.2.0 -*

Si  $\mathfrak{R} = (R_k, \zeta_k)_{k \leq K}$  une famille de référence et  $q \leq K$ , alors

$$V_{\mathfrak{R}}^* \leq \tilde{V}_{\mathfrak{R}} \leq \tilde{V}_{\mathfrak{R}}^q \leq \bar{V}_{\mathfrak{R}}$$

**Théorème 8.2.1 -**

Si une famille de référence satisfait l'hypothèse de structure en forêt (8.8), alors nous avons les égalités suivantes :

$$V_{\mathfrak{R}}^* = \tilde{V}_{\mathfrak{R}} \quad (8.10)$$

$$V_{\mathfrak{R}}^* = \tilde{V}_{\mathfrak{R}}^d \quad (8.11)$$

où

$$d = \max\{|Q|, Q \subset \{1, \dots, K\} : \forall k, k' \in Q, k \neq k' \Rightarrow R_k \cap R_{k'} = \emptyset\}$$

correspond au nombre maximum de régions disjointes différentes de notre famille de référence.

Dans ce cas particulier, nous sommes alors capable de retrouver la borne optimale  $V_{\mathfrak{R}}^*$  par le biais d'une borne calculable. Cette borne est alors utilisable sous ces hypothèses, qui se prête naturellement au cas d'expérience de protéomique.

**Construction de familles de références et calibration des  $\zeta$** 

Dans les problèmes ou expériences de protéomique, une structure hiérarchique se dégage naturellement des données, au travers des peptides, des protéines et éventuellement des gènes. Il est donc logique de considérer une famille de référence qui reflète cette hiérarchie.

La structure en arbre proposée par Durand et al. semble donc être une bonne hypothèse à considérer dans le cas d'une étude de protéomique. Les régions  $(R_k)_{k \leq K}$  de notre famille de référence peuvent être construites selon l'expérience spécifique et il reste maintenant à calibrer les  $(\zeta_k)_{k \leq K}$  associés à ces régions, de sorte à assurer le contrôle du JER (8.7). Une première méthode pour calculer les  $(\zeta_k)_{k \leq K}$  est basée sur l'inégalité de DKWM [7, 14] :

**Définition 8.2.11 -**

Pour chaque ensemble  $R_k$ ,  $k \leq K$ , d'hypothèses de notre structure, nous définissons :

$$\zeta_k = |R_k| \wedge \min_{t \in [0;1]} \left[ \left( \frac{C}{2(1-t)} + \left( \frac{C^2}{4(1-t)^2 + \frac{\sum_{i \in R_k} \mathbb{1}_{p_i > t}}{1-t}} \right)^{1/2} \right)^2 \right] \quad (8.12)$$

$$\text{où } C = \sqrt{\frac{1}{2} \log \left( \frac{K}{\alpha} \right)}.$$

La famille ainsi construite  $\mathfrak{R} = (R_k, \zeta_k)_{k \leq K}$ , contrôle le JER (8.7), sous certaines hypothèses énoncées ci-dessous :

**Proposition 8.2.1 -**

Soit la famille de p-valeurs  $(p_i)_{i \leq m}$  associée à  $\mathfrak{R}$ , avec  $\alpha \in [0, 1]$  tel que  $\alpha/K < 1/2$ .

On suppose également que :

- i Sous l'hypothèse nulle  $H_0$ , pour tout  $t \in [0, 1]$ ,  $\mathbb{P}_{H_0}(p_i \leq t) \leq t$
- ii La famille  $(p_i)_{i \in \mathcal{H}_0}$  est une famille de p-valeurs indépendantes, indépendante de  $(p_i)_{i \in \mathcal{H}_1}$

Alors, la famille de référence  $\mathfrak{R}$  contrôle le JER au niveau  $\alpha$ .

Une seconde méthode consiste à utiliser une procédure de rejet telle que celle de Holm-Bonferroni [11] sur les ensembles  $R_k$ , avec le niveau cible  $\frac{\alpha}{K}$ , pour obtenir les valeurs de  $(\zeta_k)_{k \leq K}$ . Nous considérons donc une famille de région  $(R_k)_{k \leq K}$ , ne dépendant pas de  $\alpha$ , ainsi qu'une procédure de rejet que l'on formalise par une fonction  $\phi$  ayant comme argument le niveau  $\lambda$  appliqué à la procédure, le sous-ensemble  $S$  d'hypothèses auquel on applique la procédure et nos observations  $\mathbf{X}$ . La fonction renvoie  $\phi(\mathbf{X}, \lambda, S)$  l'ensemble des hypothèses rejetées. Pour alléger les notations, nous écrivons  $\phi(\mathbf{X}, \lambda, S) = \phi(\lambda, S)$ .

*Définition 8.2.12 -*

Soit  $(R_k)_{k \leq K}$  une famille de région ne dépendant pas d'alpha, et  $\phi(\lambda, \cdot)$  une procédure de rejet de niveau  $\lambda \in [0 ; 1]$ . Alors, pour tout  $k \leq K$

$$\zeta_{tree,k}(\lambda) = |R_k| - |\phi(\lambda, R_k)|$$

Dans ce cas, si  $\phi_{HB}$  renvoie les régions de rejet pour la procédure de rejet de Holm-Bonferroni. La famille  $(R_k, \zeta_{tree,k}(\alpha/K))_{k \leq K}$  contrôle le JER au niveau  $\alpha$ . En effet,

$$\begin{aligned} \mathbb{P}(|R_k \cap \mathcal{H}_0| > \zeta_{tree,k}) &\leq \mathbb{P}(|R_k \cap \mathcal{H}_0| > |R_k| - \phi(\alpha/K, R_k)) \\ &\leq \mathbb{P}(|\phi(\alpha/K, R_k)| > |R_k| - |R_k \cap \mathcal{H}_0|) \\ &\leq \mathbb{P}(|\phi(\alpha/K, R_k) \cap \mathcal{H}_0| + |\phi(\alpha/K, R_k) \cap \mathcal{H}_1| > |R_k \cap \mathcal{H}_1|) \\ &\leq \mathbb{P}(|\phi(\alpha/K, R_k) \cap \mathcal{H}_0| > |R_k \cap \mathcal{H}_1| - |\phi(\alpha/K, R_k) \cap \mathcal{H}_1|) \\ &\leq \mathbb{P}(|\phi(\alpha/K, R_k) \cap \mathcal{H}_0| > 0) \\ &\leq \frac{\alpha}{K} \end{aligned}$$

et nous avons donc finalement :

$$JER(\mathfrak{R}) = \mathbb{P}\left(\bigcup_{k \leq K} \{|R_k \cap \mathcal{H}_0| > \zeta_{tree,k}\}\right) \leq \sum_{k \leq K} \mathbb{P}(|R_k \cap \mathcal{H}_0| > \zeta_{tree,k}) \leq \alpha$$

Cette méthode garantit également le contrôle du JER au niveau  $\alpha$  utilisé pour la procédure et n'exige pas d'indépendance entre les régions.

#### 8.2.4. Hybridation de bornes

Nous allons maintenant énoncer deux améliorations possibles des bornes précédentes, proposées par Durand et al. [6]. Supposons que deux familles de références  $\mathfrak{R}(\alpha) = (R_k(\alpha), \zeta_k(\alpha))_{k \leq K}$  et  $\mathfrak{R}'(\alpha) = (R'_k(\alpha), \zeta'_k(\alpha))_{k \leq K'}$  contrôlent le JER au niveau  $\alpha \in ]0 ; 1[$ , où le niveau  $\alpha$  choisit pour ces familles de références joue un rôle important dans l'hybridation des deux bornes et est donc mentionné explicitement. Si nous sélectionnons un sous-ensemble d'hypothèses  $S \subset \{1, \dots, m\}$  et un réel  $\gamma \in ]0 ; 1[$ , la fonction

$$V_\gamma^{hyb}(S) = \min\left(V_{\mathfrak{R}(\gamma\alpha)}^*(S), V_{\mathfrak{R}'((1-\gamma)\alpha)}^*(S)\right) \quad (8.13)$$

définit alors une borne post hoc de niveau  $\alpha$ .

Une autre hybridation est possible sur les  $\zeta$  dans le cas particulier où  $K = K'$  et  $\forall k \leq K, R_k = R'_k$  ne dépendent pas du niveau  $\alpha$  souhaité. Pour  $k \leq K$ , nous définissons

$$\zeta_k''(\gamma, \alpha) = \min(\zeta_k(\gamma\alpha), \zeta'_k((1-\gamma)\alpha))$$

et la nouvelle famille de référence en découlant  $\mathfrak{R}''(\gamma, \alpha) = (R_k, \zeta_k''(\gamma, \alpha))_{k \leq K}$ . Cette famille contrôle, comme  $\mathfrak{R}$  et  $\mathfrak{R}'$ , le JER au niveau  $\alpha$  pour tout  $\gamma \in ]0 ; 1[$ ; nous pouvons donc déduire la borne hybride

$$V_\gamma^{hyb,2}(S) = V_{\mathfrak{R}''(\gamma, \alpha)}^*(S) \quad (8.14)$$

pour un sous-ensemble  $S$  d'hypothèses. Cette borne définit une borne post hoc de niveau  $\alpha$ . Nous avons pour tout  $\alpha, \gamma \in ]0 ; 1[$  l'inégalité

$$V_{\gamma}^{hyb,2} \leq V_{\gamma}^{hyb} \quad (8.15)$$

---

## 8.3. Études de cas : applications des bornes Post-Hoc sur des données de protéomiques

---

Notre objectif est de tester la performance des différentes bornes et méthodes sur des données réelles. Nous avons commencé par travailler sur un premier jeu de données présenté en partie 8.3.1. Ce jeu de données est composé d'un mélange équimolaire de 48 protéines humaines (UPS1, Universal Proteomic Standard 1) et d'un fond de protéines d'*Arabidopsis thaliana*. Nous avons également étudié un second jeu de données, présenté dans la partie 8.3.2, qui est composé d'un mélange de 46 protéines humaines (UPS) et d'un fond de protéines de levures.

Tous les codes utilisés pour ce projet sont disponibles sur [https://github.com/aurianegbt/Post\\_Hoc](https://github.com/aurianegbt/Post_Hoc).

### 8.3.1. Un premier jeu de données

#### Présentation Générale

Le premier jeu de données utilisé est issu d'une expérience réelle et contrôlée [5], schématisée dans la figure 8.4. Pour cette expérience, un mélange équimolaire de 48 protéines humaines UPS1 (Universal Proteomic Standard 1) a été mélangé avec un fond de protéines d'*Arabidopsis thaliana*, formant ainsi 7 ensembles de 3 échantillons. Tous les ensembles ont les mêmes concentrations de protéines d'*Arabidopsis*, et une concentration croissante de protéines humaines (0.05 fmol, 0.25 fmol, 0.5 fmol, 1 fmol, 2.5 fmol, 5 fmol et 10 fmol). Les 3 ensembles avec les quantités de protéines UPS les plus faibles sont regroupés pour former une première condition appelée "low spike", et les 3 ensembles avec les quantités de protéines UPS les plus élevées sont regroupés pour former la condition "high spike". Sur les 7 conditions, nous ne conservons que les mesures de 6 d'entre elles, en supprimant celle médiane (avec 1 fmol de protéines humaines).

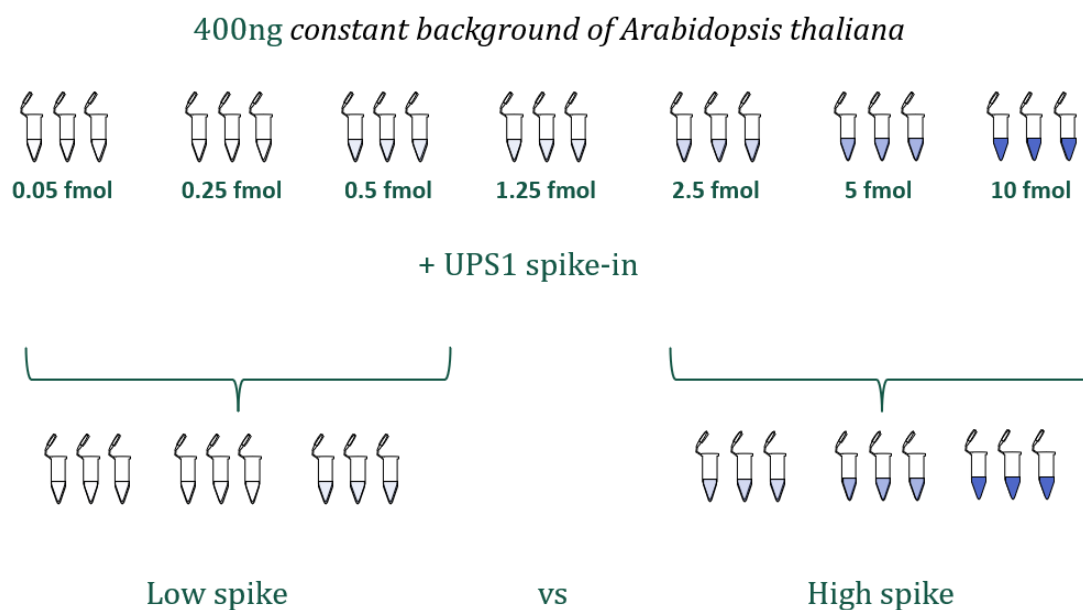


FIGURE 8.4 – Expérience effectuée pour produire le premier jeu de données étudié.

Dans chacun de ces échantillons, une analyse protéomique, comme présentée précédemment dans la partie 8.1.2, est effectuée pour mesurer l'expression des différents peptides détectés dans chaque échantillon. À partir de la littérature, et de bases de données connues sur les protéines et leurs peptides, les peptides quantifiés sont associés aux protéines qu'ils composent. Nous récupérons ainsi un jeu de données avec les mesures d'expression de 22829 peptides dans nos 18 échantillons, parmi lesquelles 21195 constituent des protéines d'*Arabidopsis* et donc 634 des protéines humaines. À chaque peptide est associé la protéine à laquelle il appartient. Nous comptabilisons 3028 protéines différentes présentes dans le jeu de données, avec exactement 2980 protéines d'*Arabidopsis* et donc nos 48 protéines humaines.

### Prétraitement des données

En protéomique, les données manquantes peuvent avoir différentes origines potentielles. Elles peuvent être causées par des limites techniques ou instrumentales lors de la collecte de données, telles que des échantillons insuffisants ou des erreurs de mesure. Une autre source vient des instruments de mesure qui ne détectent une quantité seulement au dessus d'un certain seuil ; en dessous du seuil, la quantité sera dite absente ou égale au seuil (selon le contexte). Ce problème se présente lorsque certaines protéines sont présentes à des niveaux très faibles dans l'échantillon, rendant difficile leur détection. Enfin, des problèmes de traitement et d'analyse des données peuvent également être une source de données manquantes, tels que des erreurs de prétraitement ou des problèmes de qualité de données. Il est important de prendre en compte ces différentes sources potentielles de données manquantes lors de l'analyse des données protéomiques pour éviter de fausser les résultats et assurer la fiabilité des conclusions.

En amont de l'analyse statistique, il est important de préparer les données brutes en vue de l'analyse statistique. En protéomique, les étapes préliminaires classiques sont :

- i. Filtration : on définit des règles de filtration, selon le contexte de l'expérience et les données récoltées. Cela peut être pour retirer des données manquantes, des valeurs aberrantes, etc.
- ii. Normalisation : on normalise les valeurs, ceci permet de corriger les variations de quantité de

protéines entre les échantillons, dû aux étapes d'extraction, de purification, etc. Ces variations de quantité peuvent entraîner des différences dans la quantité de protéines détectées entre les échantillons. Une fois les échantillons normalisés, les résultats des analyses seront plus précis et fiables, ce qui permettra une meilleure interprétation des résultats et une meilleure compréhension des processus biologiques étudiés [10].

- iii. Imputation : dans certains cas, on peut estimer les données manquantes, selon le contexte, pour palier au manque de données.

Ces étapes de prétraitement rendent le jeu de données fonctionnel pour y effectuer des tests. Ce prétraitement a été fait en utilisant le package DAPAR et l'application Prostar [18]. Pour l'imputation, une valeur manquante dans un set est remplacée par la moyenne des autres valeurs mesurées du même set. Cette imputation est cohérente avec la réalisation de l'expérience, où chaque échantillon d'un même set contient exactement les mêmes concentrations. Avec cette règle d'imputation, nous sommes au moins contraints de filtrer les lignes ayant un set complet manquant, et c'est donc la filtration minimale que nous avons testée. Nous avons essayé de rajouter à cette règle les suivantes :

1. **Filtration par sets plus stricte** : nous définissons, pour un peptide, ou ligne, un set comme incomplet lorsque sur ses 3 échantillons, 2 sont manquants ; et nous décidons de filtrer tous les peptides ayant des sets incomplets.
2. **Filtration avec un seuil** : nous fixons un seuil  $s \in [0 ; 1]$  et nous ne conservons les peptides, ou lignes, ayant au moins une proportion  $s$  de valeurs mesurées.

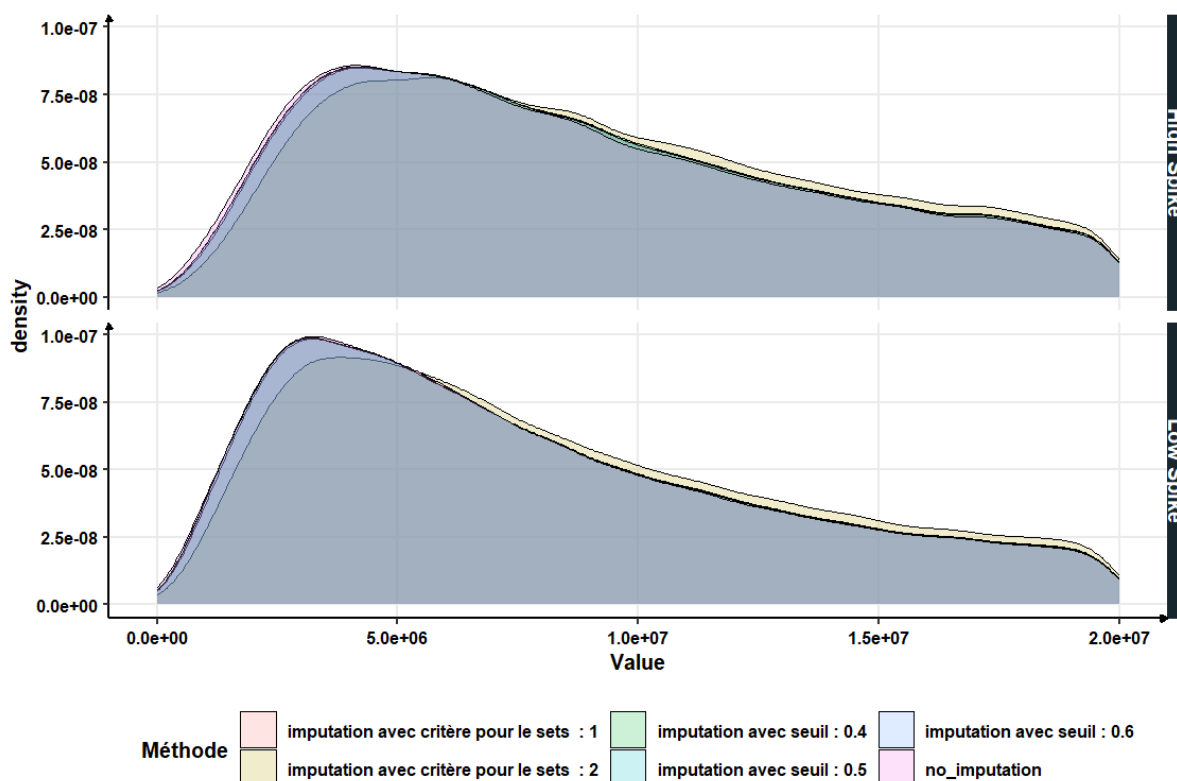


FIGURE 8.5 – Répartition des données dans chaque condition, avant et après imputation.

Nous appliquons ces méthodes à nos données, et nous visualisons le graphique des densités de nos mesures avant et après imputation sur la figure 8.5. Nous allons garder deux prétraitements parmi ceux testés. Un approchant au mieux la distribution des valeurs mesurées avant imputation, et une avec un seuil à  $s = 0.5$ .



Pour la première filtration, que nous prendrons comme référence, nous obtenons un jeu de  $m = 14427$  peptides restant, dont 14279 sont d'Arabidopsis et donc 148 sont d'humains. Notre procédure a donc filtré 8402 lignes dans notre tableau de données, pour lesquelles nous avons estimé que les données manquantes étaient gênantes pour l'étude statistique suivante. Il nous reste maintenant 2673 protéines différentes dans notre jeu de données, dont 2631 d'Arabidopsis et 42 humaines.

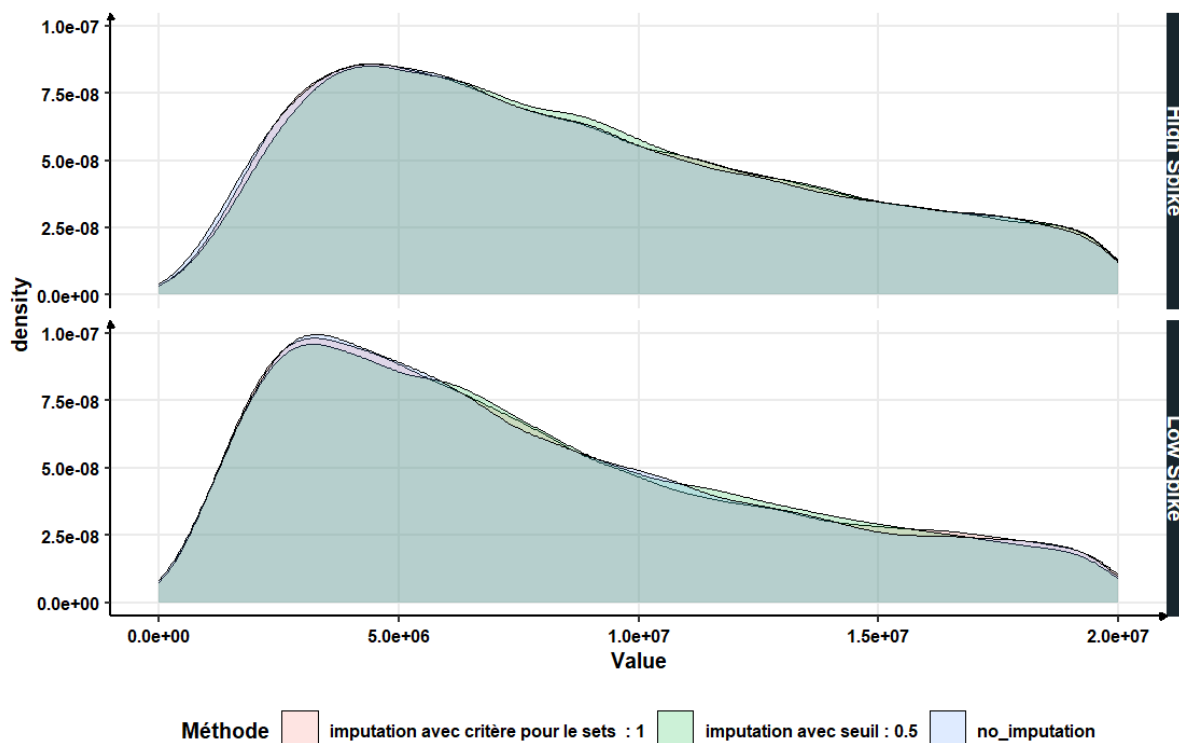


FIGURE 8.6 – Répartition des données des filtrations gardées.

Tests

Si nous reprenons la description faite précédemment, nous sommes dans le cas d'un test de deux conditions, et nous souhaitons tester si chacun de nos peptides est différentiellement abondant dans les conditions. Le test de Student, et plus précisément de Welch, présenté dans la partie 8.2.1, se prête à notre étude. Ici, la première condition est la “low spike”, et la seconde la “high spike”. Pour chaque peptide, nous allons tester si l'expression de ces peptides dans chaque condition suit la même loi. Dans notre cas, l'hypothèse  $H_0$  est “le peptide n'est pas différentiellement abondant entre les deux conditions”, et  $H_1$  l'opposé. Dans le cas de notre jeu de données naïf et contrôlé, nous savons que les hypothèses  $H_0$  correspondent aux peptides humaines. Donc nous nous attendons à observer des p-valeurs plus faibles sous  $H_1$ , et suivant une courbe suivant l'identité sur  $H_0$ . Pour effectuer ces tests, nous pouvons de nouveau utiliser l'application Prostar [18]. L'application utilise les tests du package stats [16], qui implémentent les tests bilatéraux que nous avons évoqués. Nous obtenons une p-valeur  $p_i$  pour chacun des  $i \leq m$  peptides testés. Ces p-valeurs semblent décrire une Fonction de répartition empirique des p-valeurs sous  $H_0$  trop proche de 0. Nous avons testé notre seconde méthode de prétraitement sur les données, avec le seuil. Nous obtenons la même tendance, représentée sur la figure 8.7 et 8.8.

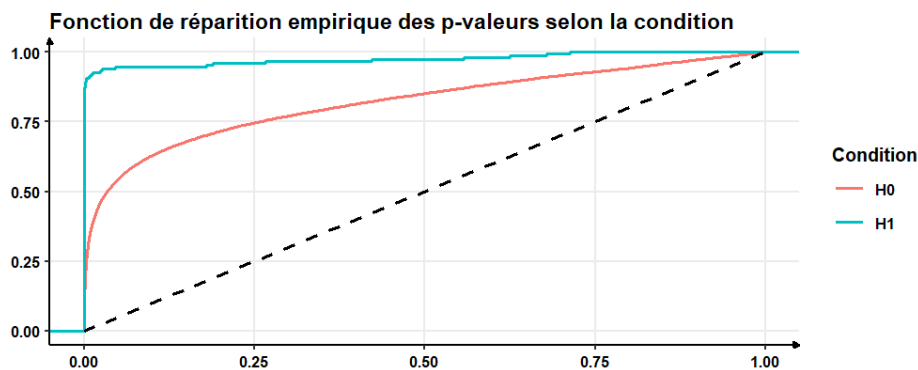


FIGURE 8.7 – Fonction de répartition empirique des p-valeurs selon  $H_0$  ou  $H_1$  avec la méthode sans seuil.

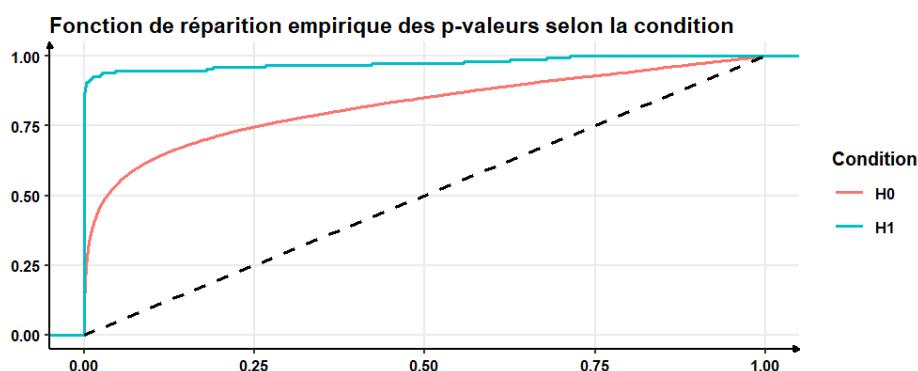


FIGURE 8.8 – Fonction de répartition empirique des p-valeurs selon  $H_0$  ou  $H_1$  avec la méthode avec seuil.

Le calcul des bornes post hoc présenté est également incohérent. Comme nous l'observons sur la figure 8.9, les différentes bornes calculées surestiment largement le nombre de positifs. Nous avons représentés sur le graphique l'oracle (TP) des vrais positifs, et les bornes  $V_{simes}$ ,  $V^*$  pour les  $\zeta_k$  obtenus par DKWM et par Holm-Bonferroni. On représente sur le graphe les fonctions  $S_k \subset \{1, \dots, m\} : |S_k| - V(S_k)$  pour  $V$  une borne post hoc, et  $S_k$  correspond à l'ensemble des hypothèses correspondant aux  $k$  plus petites p-valeur. Ces fonctions devraient donc être des bornes inférieures sur le nombre de positifs dans les données. Or ici, toutes les bornes calculées surestiment très largement le nombre de signaux dans nos données.

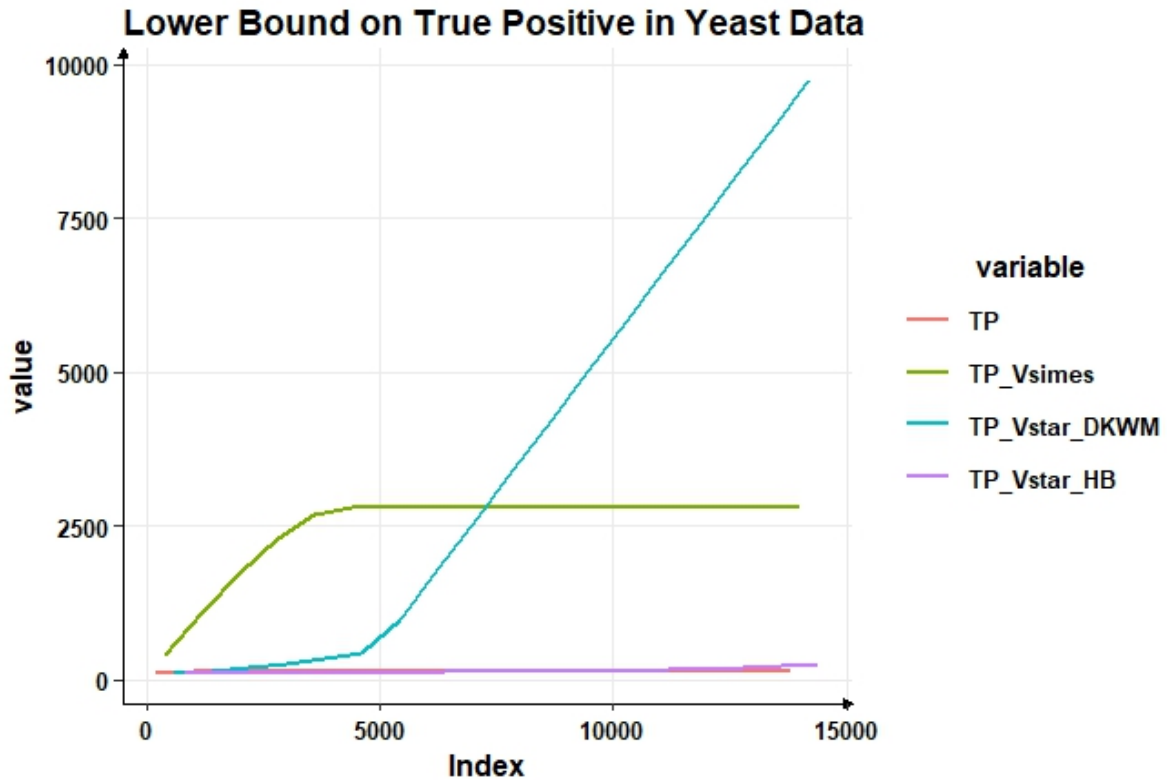


FIGURE 8.9 – Calculs des différentes bornes post hoc pour le premier jeu de données.

Ces premières observations semblent provenir de problèmes expérimentaux, tels que des défauts dans la préparation des échantillons. Le nombre élevé de faux positifs présents dans les travaux de Chion et al. [5] semblent alimenter cette hypothèse.

### 8.3.2. Étude de cas sur un jeu de données contrôlées

#### Présentation Générale et Tests

Nous considérons une seconde expérience de protéomique quantitative maîtrisée, décrite par Chion et al. [5] et réalisée de manière analogue à l'expérience précédente. On constitue 6 sets avec un fond de levure et d'une quantité différente de protéines humaines (0.5fmol, 1fmol, 10fmol, 2.5fmol, 5fmol et 25fmol). Les échantillons de 10fmol et 2.5fmol ont été inversés dans le tableau des données, et donc pour le reste des applications. On a de nouveau 6 sets de 3 échantillons identiques, qu'on associe à deux conditions "low spike" et "high spike". On effectue dans chaque échantillon une analyse, et on mesure l'expression des peptides présents dans les solutions. Notre jeu de données est composé de 21557 peptides, dont 585 humaines, pour un total de 2483 protéines, dont 46 humaines.

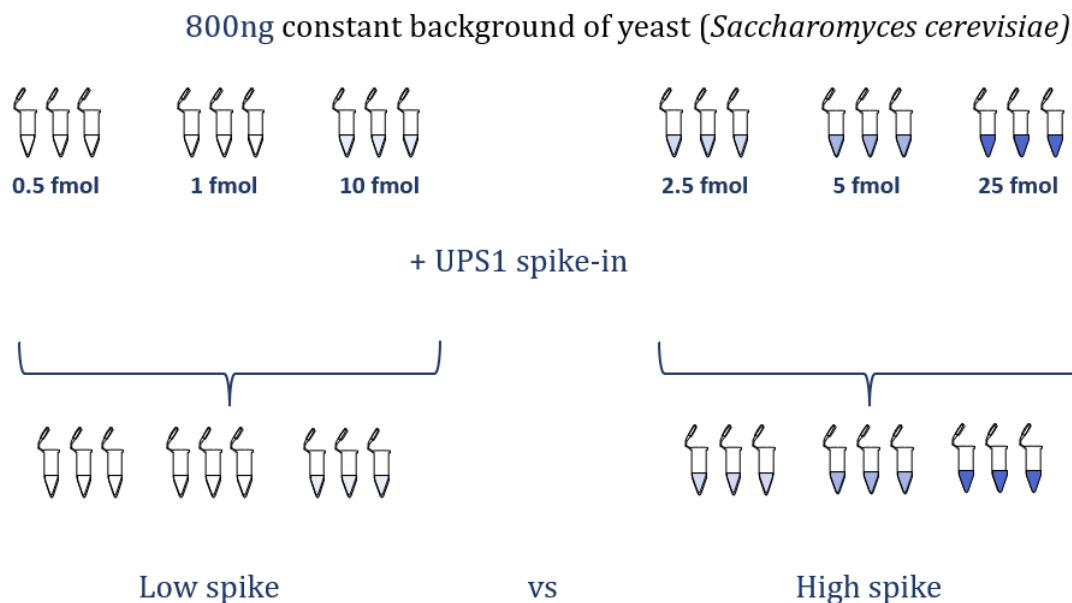


FIGURE 8.10 – Expérience effectuée pour produire le second jeu de données étudié.

Nous testons les mêmes méthodes pour le prétraitement, représentées dans la figure 8.11, et nous décidons une nouvelle fois de conserver la méthode respectant le plus la répartition avant imputation des valeurs mesurées, i.e. avec un filtre sur les lignes ayant des sets complets de valeurs manquantes. Notre filtration réduit le jeu de données à 18698 peptides, dont 192 issus de 42 protéines humaines différentes.

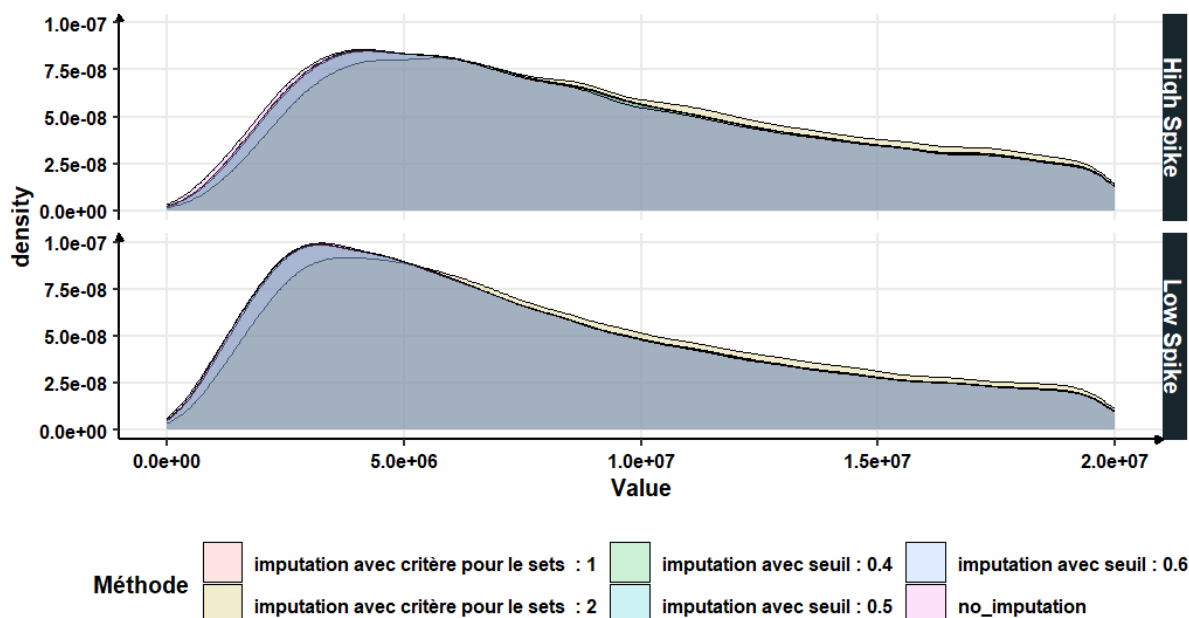


FIGURE 8.11 – Répartition des valeurs mesurées selon plusieurs méthodes de prétraitement.

Pour chacun des peptides mesurés, nous effectuons un test de Welch de niveau  $\alpha = 0.05$ , et nous

obtiennent la Fonction de répartition empirique des p-valeurs représentée en figure 8.12 . Les résultats sont cohérents avec la théorie et nos attentes. La fonction de répartition empirique des p-valeurs sous  $H_1$  est majoritairement proche de 0 et celle sous  $H_0$  décrit une courbe proche de l'identité sur  $[0; 1]$ , ce qui est plus conforme à nos attentes que les tendances observées pour l'expérience avec les protéines d'Arabidopsis. Nous remarquons également que les plus petites p-valeurs sont associées à des hypothèses de  $\mathcal{H}_0$ , ce qui est surprenant, mais probablement dû à l'inversion des échantillons de 10fmol et 2.5fmol dans la formation des conditions.

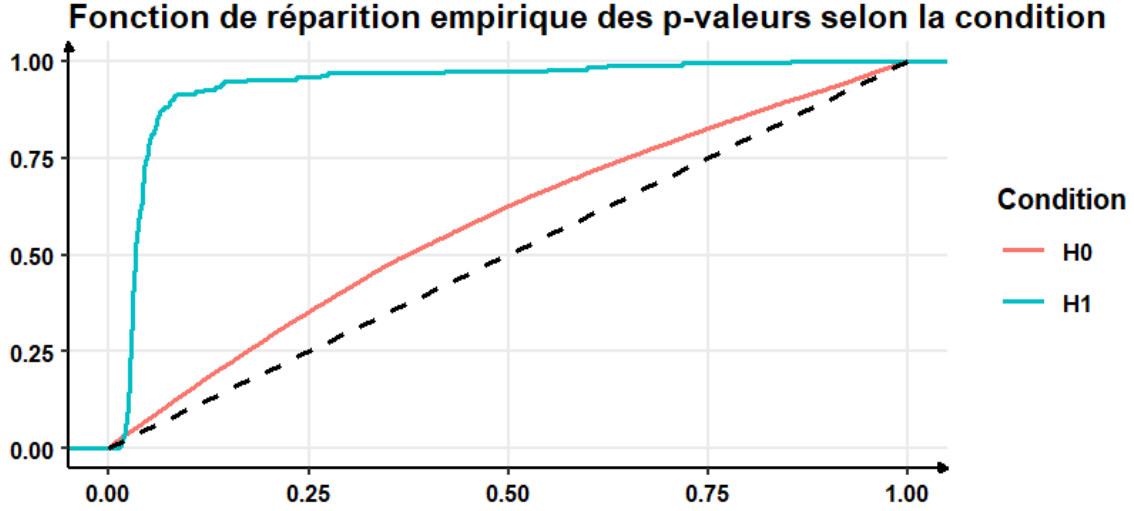


FIGURE 8.12 – Fonction de répartition empirique des p-valeurs selon la condition

### Bornes post hoc

Nous souhaitons tester les bornes post hoc présentées dans la partie 8.2. Les données se prêtent à l'hypothèse de hiérarchie proposée par Durand et al.. Chaque peptide est associé à une protéine elle même associée à une espèce parmi levure et humaine. On construit donc la famille de référence notée  $\mathfrak{R} = (R_k, \zeta_k)_{k \leq K}$  pour nos données. En particulier, ici,  $R_1 = \{1, \dots, m\}$ ,  $R_2 = \{1, \dots, n_{\text{yeast}}\}$  et  $R_3 = \{n_{\text{yeast}}+1, \dots, m\}$  où  $n_{\text{yeast}} = 18506$  correspond au nombre de peptides de levure dans nos données. Nous avons donc testé les bornes  $V_{\text{simes}}$  (Définition 8.2.2)  $V_{\mathfrak{R}}^*$  (Définition 8.2.2) pour deux constructions de  $(\zeta_k)_{k \leq K}$  évoquées : une par l'inégalité DKWM (Définition 8.2.3), et une par procédure de Holm-Bonferroni. Ces deux dernières bornes sont respectivement notées  $V_{DKWM}^*$  et  $V_{HB}^*$  pour la suite.

La figure 8.13 représente les fonctions  $k \in \{1, \dots, m\} \mapsto |S_k| - V(S_k)$ , où  $V$  est une des bornes post hoc citées et  $S_k$  est l'ensemble des  $k$  hypothèses ayant les  $k$  plus petites p-valeurs. La figure 8.14 est une restriction de la figure 8.13 aux individus tels que  $k \leq 2000$ . Les courbes tracées correspondent donc à une borne inférieure sur le nombre de vrais positifs dans l'ensemble considéré ; nous avons donc tracé la courbe du nombre de vrais positifs dans chacun des ensembles  $S_k$ , notée TP. Nous remarquons alors que la courbe  $V_{DKWM}^*$  est bien meilleure que les autres bornes, dans le sens où c'est elle qui fournit la borne inférieure sur le nombre de vrais positifs la plus fine. La courbe semble suivre la vraie tendance, mais explose vers  $k = 14000$ . Nous pensons à ce stade que la dépendance, exigé par DKWM, dans les données peut être à l'origine de l'explosion à partir de  $k = 14000$ . La courbe  $V_{\text{simes}}$  reste supérieure à celle  $V_{HB}^*$  pour une majorité d'ensemble  $S_k$  ; mais pour  $k$  grand,  $k \geq 16500$ , la tendance s'inverse, et la borne  $V_{HB}^*$  est capable de tirer plus d'information de la totalité des résultats. Ces deux dernières courbes restent

bien inférieures pour n'importe lequel de ces ensembles  $S_k$  à TP, et donc surestime le nombre de faux positifs dans les hypothèses sélectionnées.

Contrairement à la borne  $V_{simes}$ ,  $V_{HB}^*$  et  $V_{DKWM}^*$  ne considère pas les plus petites p-valeurs comme du signal et ne fait donc pas d'erreur sur ces parties. La procédure post hoc, avec l'hypothèse de hiérarchie, permettrait donc d'éviter des faux positifs sur les hypothèses de p-valeurs faussement significatives.

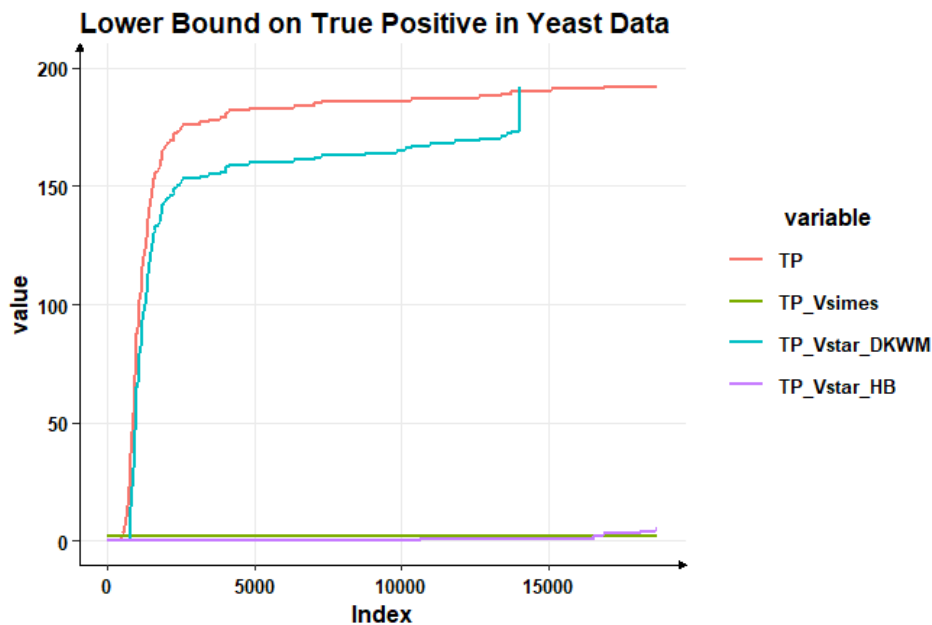


FIGURE 8.13 – Calcul des différentes bornes post hoc sur le second jeu de données.

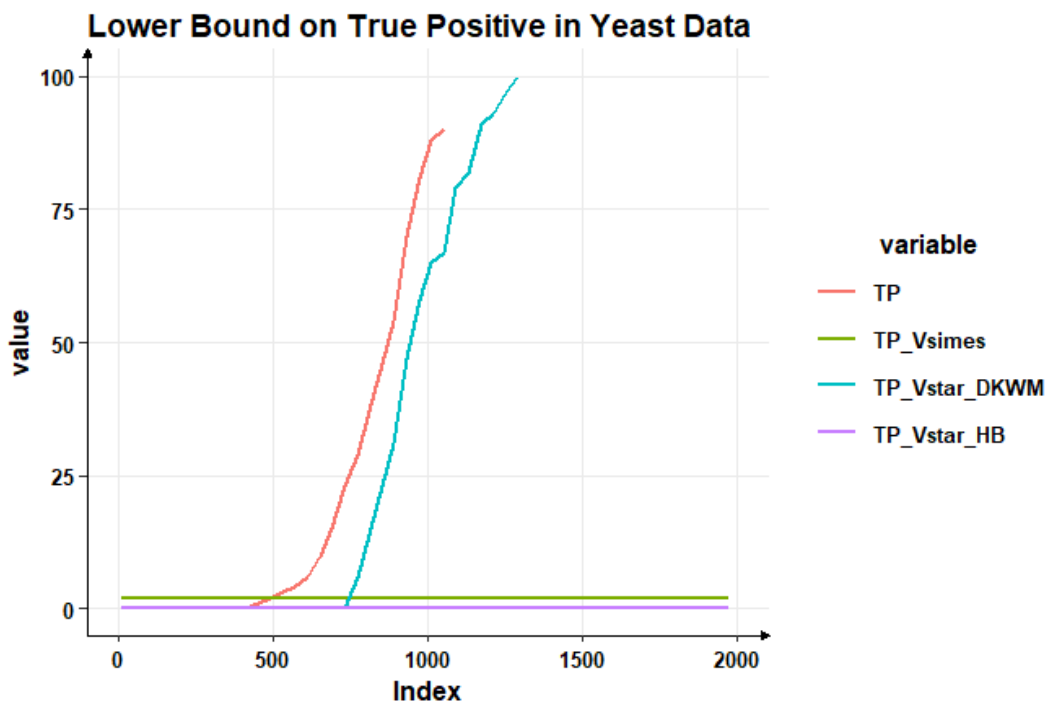


FIGURE 8.14 – Calcul des différentes bornes post hoc sur le second jeu de données pour les plus petites p-valeurs.

### Test de la $\lambda$ -calibration

Au vu des résultats précédents, nous avons testé la méthode de  $\lambda$ -calibration, présentée dans la partie 8.2.2. Pour cela, nous utilisons la méthode implémenté directement dans le package sanssouci [15]. Les résultats obtenus sont sur la figure 8.15. La courbe obtenue par la méthode de  $\lambda$ -calibration, noté  $V_{simes}^{cal}$  est globalement meilleure que les 2 bornes  $V_{simes}$  et  $V_{HB}^*$ , mais commet une fausse estimation du nombre de faux positifs sur les plus petites p-valeurs. Etant donné que la méthode de  $\lambda$ -calibration est valide sous une quelconque forme de dépendance, cela nous renforce à penser que les plus petites p-valeurs viennent de la mauvaise attribution des échantillons dans les conditions. Hormis pour les ensembles  $k \geq 14000$ , la borne  $V_{DKWM}^*$  semble fournir une bonne borne pour le nombre de positifs dans un échantillon. Dans la pratique, cette explosion pourrait ne pas poser de problème puisqu'il est manifeste, même sans connaissance sur  $\mathcal{H}_0$  que le calcul de la borne est erronée pour ces grandes valeurs, puisque la courbe devient linéaire, ce qui est absurde par rapport à la théorie où on s'attendrait à avoir une courbe qui s'applatit sur les plus grandes p-valeurs. L'explosion de cette courbe serait alors remarquée dans tous les cas, et seule le début de la courbe serait alors exploitée. Si l'on vérifie la structure  $(R_k, \zeta_k)$  générée, le JER est violé pour quelques protéines de levure sur les feuilles mais globalement pour l'ensemble des protéines de levure, c'est-à-dire que pour  $R_2 = \{ \text{peptides de levures} \}$  complet le JER est violé. Donc même si nous retirions seulement les feuilles violons le JER, le contrôle ne serait toujours pas vérifié sur la globalité des peptides.

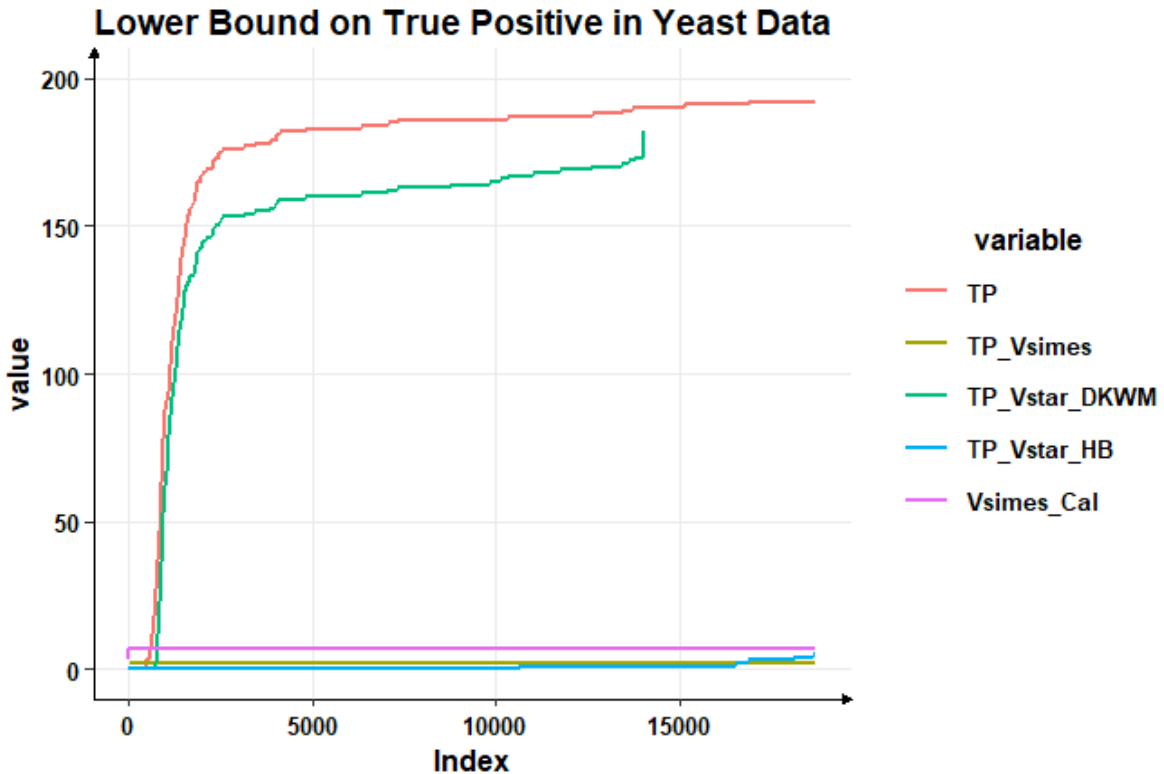


FIGURE 8.15 – Calcul des différentes bornes, et de bornes par la méthode de  $\lambda$ -calibration, sur le second jeu de données.

### Hybridation des bornes

Nous avons également implémenté les différentes versions d'hybridation proposées dans la partie 8.2.4. Les deux bornes hybrides calculées (8.13), (8.14) ont été testées pour  $\gamma \in \{0.1, 0.2, \dots, 0.9\}$  et toujours  $\alpha = 0.05$ , sur les ensembles  $S_k \subset \{1, \dots, m\}$  des hypothèses correspondant aux  $k$  plus petites p-valeurs. On considère comme familles de références ici les familles  $(R_k, \zeta_{DKWM,k}(\gamma\alpha))_{k \leq K}$  et  $(R_k, \zeta_{HB,k}((1-\gamma)\alpha))_{k \leq K}$  où les ensembles  $(R_k)_{k \leq K}$  sont ceux présentés précédemment et issus de la hiérarchie.

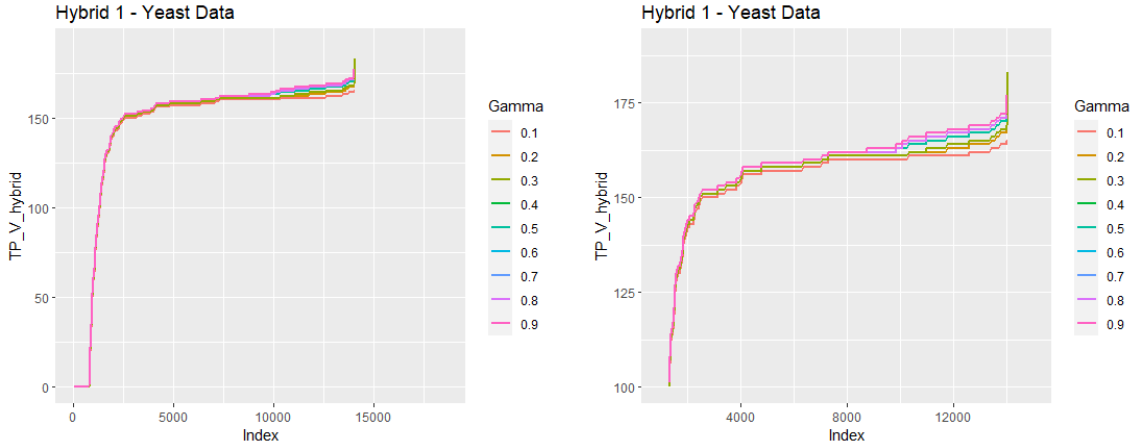


FIGURE 8.16 – Calcul de  $V_{\gamma}^{hyb}$

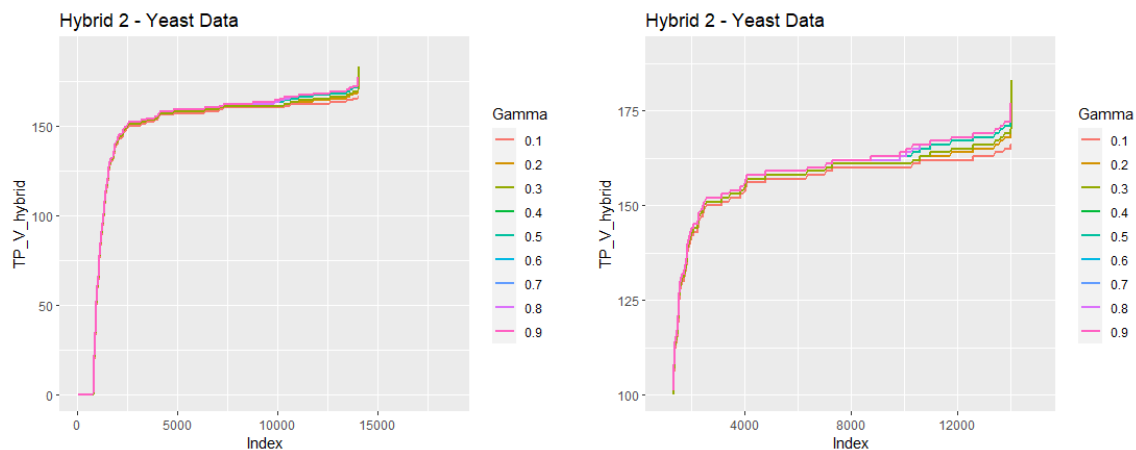
Les résultats obtenus sont illustrés en figures 8.16 et 8.17. Sur l'ensemble des 2500 plus petites p-valeurs, toutes les bornes  $V_{\gamma}^{hyb}$  sont égales, pour toute valeur de  $\gamma$ . La différence entre les courbes  $V_{\gamma}^{hyb}$  n'est visible que pour des p-valeurs assez grandes ( $k \geq 10000$ ), là les bornes restent superposées puis se séparent en différentes courbes pour chaque valeur de  $\gamma$ . On a l'inégalité :

$$V_{\gamma}^{hyb} \leq V_{\gamma'}^{hyb}, \quad \forall \gamma \leq \gamma'$$

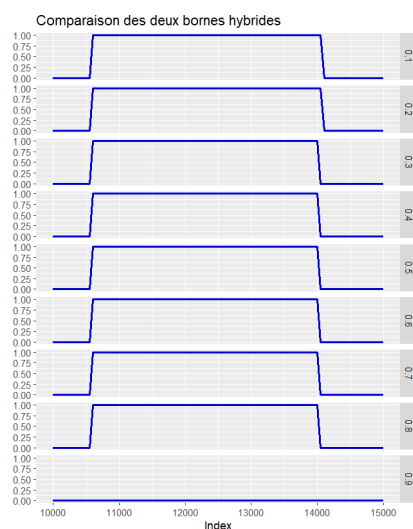
Il semble donc optimal de prendre un  $\gamma$  plus élevé et donc de diminuer le niveau de la borne  $V_{DKWM}^*$ .

La seconde borne  $V_{\gamma}^{hyb,2}$  se comporte de la même manière, les différentes bornes sont égales pour les 2500 plus petites p-valeurs, puis se sépare dans le même ordre. La borne la plus performante étant donc celle pour le  $\gamma$  le plus élevé (ici  $\gamma = 0.9$ ).



FIGURE 8.17 – Calcul de  $V_\gamma^{hyb,2}$ 

Nous avons énoncé l'inégalité (8.15) précédemment, et nous avons donc visualisé, figure 8.18, la différence entre les deux bornes  $V_\gamma^{hyb,2} - V_\gamma^{hyb}$ . On peut alors constater que l'inégalité peut être stricte et donc les bornes  $V_\gamma^{hyb,2}$  et  $V_\gamma^{hyb}$  ne sont pas équivalentes. La borne  $V_\gamma^{hyb,2}$  est une amélioration de la seconde. Dans le cas présent, la seconde borne est sensiblement meilleure sur certains ensembles, et les deux bornes restent équivalentes dans le cas optimale, i.e.  $\gamma \in [0; 1]$  grand

FIGURE 8.18 – Comparaison des deux bornes  $V_\gamma^{hyb,2}$  et  $V_\gamma^{hyb}$ 

### Biais de sélection

Comme nous avons énoncé en introduction, il est fréquent qu'en biologie les analyses statistiques soient effectués sur un sous-ensemble d'hypothèses, dont les p-valeurs sont par exemple les  $k$  plus petites. Ceci induit un biais dans la sélection qui compromet les garanties statistiques usuelles. Néanmoins, les bornes post hoc sont résistantes à ce biais, et nous avons donc décidé de regarder les bornes obtenues pour ces ensembles sélectionnés, afin de vérifier la tenue de ces garanties. Nous avons donc considéré les ensembles  $S_k$  des hypothèses rejetées par les différentes procédures de rejet appliquées au  $k$  plus petites p-valeurs, pour différentes valeurs de  $k \in \mathcal{K} = \{500 \times n, 1 \leq n \leq 10\}$ . Avec le formalisme introduit dans le

paragraphe 8.2.3, nous avons :

$$S_k = \phi \left( \alpha, \{i \in \{1, \dots, m\} : (p_i) \in (p_{(s)})_{1 \leq s \leq k}\} \right)$$

Pour chacun de ces ensembles, nous calculons la valeur de  $V_{DKWM}^*(S_k)$ . Les résultats sont présentés en figure 8.19, mais résumés dans la table 8.1 pour plus de lisibilité.

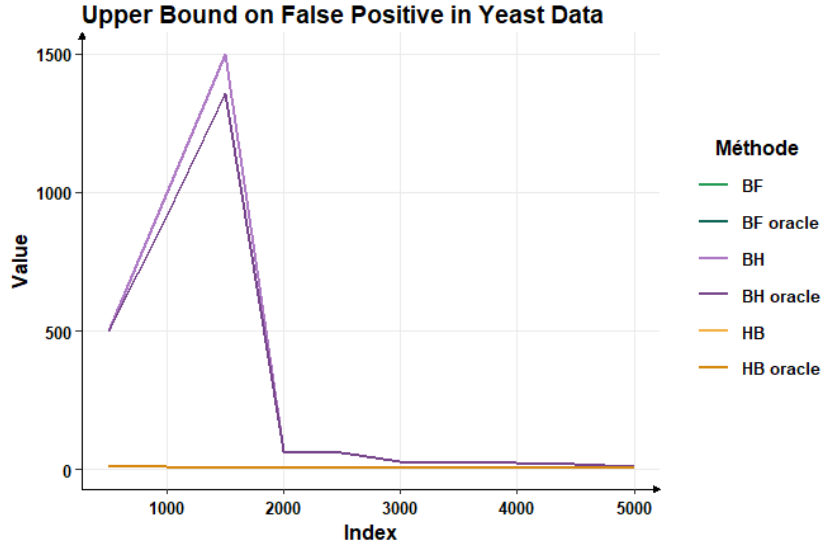


FIGURE 8.19 – Application de  $V_{DKWM}^*$  dans le cas d’une sélection dans les données, les méthode ”oracle“ correspondent au vrai nombre de faux positifs obtenus sur l’ensemble  $S_k$  obtenu par la méthode de rejet mentionné .

Nous avons donc testé chacune des procédures présentées en partie 8.2.1. Les procédures de Bonferroni, Holm-Bonferroni et Benjamini-Hochberg sont donc appliquées à l’ensemble des  $k$  plus petites p-valeurs obtenues, donnant un ensemble  $S_k$ . Pour les premières procédures, il y a égalité entre  $|S_k|$ ,  $V_{DKWM}^*(S_k)$  et  $|S_k \cap \mathcal{H}_1|$  (à l’exception de  $k = 1000$  pour Holm-Bonferroni et  $k \in \{500, 1000, 1500\}$ ). Pour la dernière procédure, la borne  $V_{DKWM}^*$  est exactement égale dans la majorité des cas au nombre de faux positifs dans l’ensemble rejeté  $S_k$ , et est donc une borne précise.

Nous avons remarqué précédemment que les plus petites p-valeurs obtenues correspondaient à des faux positifs, puisqu’il s’agissait de peptides de levures. L’oracle  $|S_k \cap \mathcal{H}_1|$  est égale dans la majorité des cas au cardinal de l’ensemble rejeté, donc on se trompe systématiquement lors des procédures de rejets d’une part, mais la borne obtenue pour  $V_{DKWM}^*$  est correcte mais égale au cardinal de l’ensemble rejeté, donc en pratique on ne peut pas tirer d’informations sur ces ensembles, ce qui confirme que pré-sélectionner les données n’est pas une chose à faire, et on ne conserve pas les garanties théoriques. Néanmoins, à mesure que  $k$  augmente, nous avons d’autant plus d’informations sur les données et le nombre d’hypothèses rejetées diminuent. Plus l’expérience est testée sur un nombre  $k$  important de données, moins les bornes et procédures ne se font piéger par les plus petites p-valeurs erronées. Il semble donc d’autant plus pertinent de ne pas pré-sélectionner en amont et bien d’exploiter la totalité des informations disponibles.

| Procédure utilisée | Valeur de k | $ S_k $ | $V_{DKWM}^*(S_k)$ | $ S_k \cap \mathcal{H}_0 $ | FDP( $S_k$ ) |
|--------------------|-------------|---------|-------------------|----------------------------|--------------|
| Bonferroni         | 500         | 10      | 10                | 10                         | 1            |
|                    | 1000        | 8       | 8                 | 8                          | 1            |
|                    | 1500        | 5       | 5                 | 5                          | 1            |
|                    | 2000        | 5       | 5                 | 5                          | 1            |
|                    | 2500        | 5       | 5                 | 5                          | 1            |
|                    | 3000        | 5       | 5                 | 5                          | 1            |
|                    | 3500        | 5       | 5                 | 5                          | 1            |
|                    | 4000        | 5       | 5                 | 5                          | 1            |
|                    | 4500        | 5       | 5                 | 5                          | 1            |
|                    | 5000        | 4       | 4                 | 4                          | 1            |
| Holm-Bonferroni    | 500         | 13      | 13                | 13                         | 1            |
|                    | 1000        | 9       | 9                 | 8                          | 1            |
|                    | 1500        | 6       | 6                 | 6                          | 1            |
|                    | 2000        | 5       | 5                 | 5                          | 1            |
|                    | 2500        | 5       | 5                 | 5                          | 1            |
|                    | 3000        | 5       | 5                 | 5                          | 1            |
|                    | 3500        | 5       | 5                 | 5                          | 1            |
|                    | 4000        | 5       | 5                 | 5                          | 1            |
|                    | 4500        | 5       | 5                 | 5                          | 1            |
|                    | 5000        | 4       | 4                 | 4                          | 1            |
| Benjamini-Hochberg | 500         | 500     | 500               | 498                        | 0.996        |
|                    | 1000        | 1000    | 1000              | 915                        | 0.915        |
|                    | 1500        | 1500    | 1500              | 1355                       | 0.903        |
|                    | 2000        | 66      | 66                | 66                         | 1            |
|                    | 2500        | 59      | 59                | 59                         | 1            |
|                    | 3000        | 26      | 26                | 26                         | 1            |
|                    | 3500        | 24      | 24                | 24                         | 1            |
|                    | 4000        | 23      | 23                | 23                         | 1            |
|                    | 4500        | 17      | 17                | 17                         | 1            |
|                    | 5000        | 10      | 10                | 10                         | 1            |

TABLE 8.1 –  $V_{DKWM}^*$  calculé sur  $(S_k)_{k \in \mathcal{K}}$ , correspondant au résultat des procédures de rejet appliquées aux  $k$  plus petites p-valeurs obtenues, l'oracle correspond au vrai nombre de faux positifs  $|S_k \cap \mathcal{H}_1|$  dans cet ensemble  $S_k$ , et le False Discovery Proportion (FDP) de  $S_k$ .

---

## Conclusion

---

Pour ce projet, nous avons donc pu appliquer les bornes Post-Hoc développées par G. DURAND, G. BLANCHARD, P. NEUVIAL et E. ROQUAIN à deux jeux de données réels. Tous les codes utilisés pour ce projet sont disponibles sur [https://github.com/aurianegbt/Post\\_Hoc](https://github.com/aurianegbt/Post_Hoc).

Nous avons pu constater l'efficacité et la fiabilité dans certains cas de la borne  $V_{DKWM}^*$ . En effet, les résultats absurdes obtenus pour les plus grandes p-valeurs sont remarquables sans connaissance à priori sur les données, et donc correctement interprétables. Sans cette explosion, et dès lors que les sélections sont généralement effectuées sur des ensembles de plus petites p-valeurs, cette borne est efficace sur les données et par l'hypothèse de hiérarchie, ne commet pas d'erreur sur le premier palier des plus petites p-valeur pourtant faussement significatives. Néanmoins, les problèmes rencontrés suggèrent que les bornes et résultats sont sensibles à la dépendance dans les données. Il est nécessaire de développer de nouvelles bornes adaptées à cette dépendance.

Nous avons pu observer également que la borne était toujours efficace dans le cas de biais de sélection décrit précédemment et en introduction, ce qui en fait un outil statistique fiable et robuste en protéomique. Les versions hybrides de nos bornes, testées sur données réelles, ont montré que l'inégalité (8.15) pouvait être stricte et donc que la borne  $V_{\gamma}^{hyb,2}$  pouvait être meilleure lorsqu'elle était applicable, et le choix d'un  $\gamma$  plus grand donnait également de meilleurs résultats. La seconde borne est donc une amélioration de  $V^{hyb}$  et non pas seulement une borne équivalente.

Nous avons travaillé sur deux jeux de données réels et maîtrisés. Ces données nous permettent donc d'avancer vers un contexte réaliste de leurs applications possibles, et montre une première fois que ces procédures Post-Hoc fournissent des outils statistiques robustes aux problèmes de sélection rencontrés. Ce travail ouvre donc de nouvelles perspectives pour l'analyse de données issues d'expériences réelles dans lesquelles la composition des échantillons n'est pas connue en avance.

---

## Liste des références

---

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate : A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57(1) :289–300.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29(4) :1165 – 1188.

- Blanchard, G., Neuvial, P., and Roquain, E. (2020). Post hoc confidence bounds on false positives using reference families. *The Annals of Statistics*, 48(3) :1281 – 1303.
- Bonferroni, C. E. (1935). Il calcolo delle assicurazioni su gruppi di teste.
- Chion, M., Carapito, C., and Bertrand, F. (2022). Accounting for multiple imputation-induced variability for differential analysis in mass spectrometry-based label-free quantitative proteomics. *PLOS Computational Biology*, 18(8) :1–26.
- Durand, G., Blanchard, G., Neuvial, P., and Roquain, E. (2020). Post hoc false positive control for structured hypotheses. *Scandinavian Journal of Statistics*, 47(4) :1114–1148.
- Dvoretzky, A., Kiefer, J., and Wolfowitz, J. (1956). Asymptotic Minimax Character of the Sample Distribution Function and of the Classical Multinomial Estimator. *The Annals of Mathematical Statistics*, 27(3) :642 – 669.
- Goeman, J. J. and Solari, A. (2011). Multiple Testing for Exploratory Research. *Statistical Science*, 26(4) :584 – 597.
- Gosset, W. S. (1908). The probable error of a mean. *Biometrika*, 6(1) :1–25.
- Graw, S., Tang, J., Zafar, M. K., Byrd, A. K., Bolden, C., Peterson, E. C., and Byrum, S. D. (2020). proteinorm – a user-friendly tool for normalization and analysis of tmt and label-free protein quantification. *ACS Omega*, 5(40) :25625–25633. PMID : 33073088.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6 :65–70.
- James, G. S. (2018). The Behrens-Fisher Distribution and Weighted Means. *Journal of the Royal Statistical Society : Series B (Methodological)*, 21(1) :73–90.
- Lualdi, M. and Fasano, M. (2019). Statistical analysis of proteomics data : A review on feature selection. *Journal of Proteomics*, 198 :18–26. 10 Year Anniversary of Proteomics.
- Massart, P. (1990). The Tight Constant in the Dvoretzky-Kiefer-Wolfowitz Inequality. *The Annals of Probability*, 18(3) :1269 – 1283.
- Neuvial, P., Blanchard, G., Durand, G., Enjalbert-Courrech, N., and Roquain, E. (2022). *sanssouci : Post Hoc Multiple Testing Inference*. R package version 0.12.8.
- R Core Team (2013). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Welch, B. L. (1947). The generalization of ‘student’s’ problem when several different population variances are involved. *Biometrika*, 34(1/2) :28–35.
- Wieczorek, S., Combes, F., Lazar, C., Gai Gianetto, Q., Gatto, L., Dorffer, A., Hesse, A.-M., Couté, Y., Ferro, M., Bruley, C., and Burger, T. (2016). DAPAR, ProStaR : software to perform statistical analyses in quantitative discovery proteomics. *Bioinformatics*, 33(1) :135–136.

# Stratégie de construction de modèle



Auriane GABAUT  
Encadré par Mélanie PRAGUE  
Avril 2023 - Septembre 2023

## Sommaire

|                                                                        |            |
|------------------------------------------------------------------------|------------|
| <b>Introduction et Motivations</b>                                     | <b>178</b> |
| <b>9.1 Préliminaires</b>                                               | <b>179</b> |
| 9.1.1 Modèles non-linéaire à effets mixtes                             | 180        |
| 9.1.2 Algorithmes de construction de modèle                            | 187        |
| <b>9.2 Sélection du modèle de covariables par LASSO dans SAMBA</b>     | <b>191</b> |
| 9.2.1 Least Absolute Shrinkage and Selection Operator (LASSO)          | 191        |
| 9.2.2 Implémentation du LASSO pour la sélection de variable dans SAMBA | 193        |
| 9.2.3 Décroissance du critère de sélection au cours des itération      | 195        |
| <b>9.3 Application à la dynamique hummorale post-boost</b>             | <b>196</b> |
| 9.3.1 Cadre d'application                                              | 196        |
| 9.3.2 Application à des simulations                                    | 198        |
| 9.3.3 Prevac-UP                                                        | 205        |

## Introduction et Motivations

La maladie d'Ébola est une infection virale grave et souvent mortelle qui a suscité de vives inquiétudes au niveau mondial en raison de ses épidémies dévastatrices. L'épidémie, par exemple, ayant eu lieu en Guinée, en Sierra Leone et au Liberia de janvier 2015 à juin 2016, a provoqué près de 30 000 cas et de 11 310 décès. Au vu de ces bilans dramatiques, il devient évident que la recherche médicale, en vaccinologie, est d'une importance cruciale. Une des approches essentielles pour lutter contre cette maladie est la vaccination, les essais ont notamment montré une efficacité supérieure à 75 % contre l'infection [5].

De fait, il est donc crucial de mieux comprendre les mécanismes sous-jacents de la réponse immunitaire afin de mettre en place des stratégies vaccinales efficaces pour un grand nombre de personnes.

C'est dans ce contexte que l'équipe SISTM de l'INRIA de Bordeaux, *Statistics In System biology and Translational Medicine*, travaille. Spécialisée dans l'analyse de données de grande dimension, l'équipe se concentre sur l'utilisation de méthodes statistiques adaptées pour aborder des questions cliniques et biologiques. Leur recherche est notamment orientée sur la modélisation mécanistique et les méthodes d'apprentissage statistique. Leur travail est principalement consacré aux applications dans la recherche clinique, avec une emphase particulière sur l'immunologie et la vaccinologie.

Un exemple concret de problématique rencontrée est le choix du modèle pour modéliser au mieux la dynamique de la réponse humorale à un vaccin, en sélectionnant notamment les covariables qui influent significativement sur le système parmi ceux observés ou pouvant être collectés. Étant donné que chaque individu est caractérisé par un ensemble de variables, appelées covariables, il revêt une importance cruciale de déterminer lesquelles de ces variables ont une influence significative sur le phénomène biologique, et grâce auxquelles nous pourrions prédire efficacement la réponse à un vaccin. En identifiant précisément les variables clés, nous pouvons non seulement mieux comprendre les phénomènes biologiques, mais également renforcer notre capacité à prédire les dynamiques virales. Cette avancée a un potentiel immense pour améliorer la prévention et la gestion des épidémies, ouvrant ainsi la voie à des interventions plus efficaces et ciblées dans la lutte contre des maladies graves telles que la maladie d'Ébola. C'est sur cette question, de sélection de variables, que se concentre mon mémoire.

L'objectif de mon travail est de proposer et d'implémenter une méthode de sélection des covariables dans un algorithme déjà performant, mais qui n'est pas parfaitement adapté à la grande dimension [31]. Ce besoin de nouveaux outils gérant les données de grandes dimensions vient de l'explosion récente du nombre de ce type de données. Par exemple, le projet COVERAGE-Immuno, mené par l'équipe SISTM, a récemment montré qu'il était possible de récupérer les données de transcriptomique d'un individu à un temps en effectuant juste un prick test [10].

Après avoir passé en revue les méthodes existantes et les notions centrales abordées pendant ce stage, je présenterai en détail la méthode de sélection des covariables développées lors de ce stage au sein de l'équipe SISTM et comment elle se distingue des approches existantes. Les résultats d'applications sur des simulations et la méthode qui sera utilisée pour l'application à un vrai jeu de données provenant de l'étude Prevac-UP [3] est également présentée. Enfin, je discuterai des résultats obtenus et des perspectives futures en ce qui concerne cette problématique et ma thèse au sein de l'équipe.

---

## 9.1. Préliminaires

---

La dynamique des anticorps joue un rôle essentiel dans la réponse immunitaire. Comprendre les mécanismes biologiques qui contrôlent cette dynamique est donc crucial pour la recherche médicale et le développement de stratégies vaccinales efficaces. Pour étudier ces phénomènes complexes, nous pouvons utiliser des modèles qui permettent de décrire et d'expliquer les dynamiques cellulaires impliquées.

### 9.1.1. Modèles non-linéaire à effets mixtes

Une approche couramment empruntée pour modéliser ces phénomènes est l'utilisation des modèles à effets mixtes [19]. Les modèles à effets mixtes sont des outils statistiques puissants qui permettent d'analyser les données provenant d'études expérimentales complexes, prenant notamment en compte l'hétérogénéité inter- et intraindividuelle présentes au sein d'une population. Ils prennent en compte à la fois les facteurs fixes, qui influencent de manière constante les réponses immunitaires humores au sein d'une population, mais aussi les effets aléatoires, qui capturent les variations intraindividuelles. Ces modèles offrent une approche statistique robuste pour étudier les différences interindividuelles et intraindividuelles dans les réponses immunitaires humores.

#### Modèles linéaires à effets mixtes

Au cours d'un essai clinique, des données biologiques, cliniques, etc., sont collectées au cours du temps pour un nombre  $N$  d'individus et ces données sont susceptibles de varier d'un individu à l'autre selon les différentes caractéristiques de l'individu, comme son âge, son sexe, etc, mais aussi selon une certaine variabilité résiduelle non capturée par les observations. C'est cette variation que les modèles à effets mixtes permettent de capturer. De manière générale, la variable observée s'écrit  $Y_i = (Y_{ij}) \in \mathbb{R}^{n_i}$  pour un individu  $i \leq N$  à son  $j^{\text{ème}}$  temps  $t_{ij}$  de mesure,  $j \leq n_i$ , et  $X_{ij}^T = (X_{ij}^1, \dots, X_{ij}^n)^T \in \mathbb{R}^n$  est le vecteur contenant les informations, ou covariables, sur l'individu  $i$  au temps  $t_{ij}$ . La matrice des covariables de l'individu  $i$  s'écrit alors  $X_i = (X_{i1}, \dots, X_{in_i})^T \in \mathcal{M}_{n_i n}(\mathbb{R})$ .

Finalement, le modèle linéaire usuel liant  $Y_i$ , supposé centré sans perte de généralité, à  $X_i$  pour l'individu  $i$  s'écrit

$$Y_i = X_i \beta + \varepsilon_i$$

où  $\varepsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, \Sigma_i)$ . Les paramètres du modèle sont les coefficients contenus dans le paramètre  $\beta = (\beta_1, \dots, \beta_m)^T \in \mathbb{R}^m$ , et  $\sigma^2 > 0$  propre au modèle d'erreur.

Dans le cadre des modèles à effet-mixtes, il est supposé qu'une partie ou la totalité des paramètres du modèle sont eux aussi individuels, et que donc nous définissons pour ces derniers un modèle de régression propre à chaque individu. Pour faire cela, nous définissons les paramètres comme la somme d'un paramètre de population et d'un effet aléatoire modélisant cette variabilité individuelle. Pour un paramètre  $\beta_p \in \mathbb{R}$ , nous définissons finalement le paramètre individuel comme

$$\forall i \leq N, \beta_{pi} = \beta_{p0} + \eta_{pi}$$

avec  $\eta_{pi} \stackrel{iid}{\sim} \mathcal{N}(0, \omega_p^2)$  l'effet aléatoire, et  $\beta_{p0} \in \mathbb{R}$  l'effet fixe.

De manière générale, avec les notations précédentes et si  $p_1, \dots, p_q, q \leq p$ , représentent les indices des coefficients de  $\beta$  auxquels on ajoute un effet aléatoire  $\eta_i = (\eta_{p_1}, \dots, \eta_{p_q})^T$ ; alors, pour  $Z_i$  la sous-matrice



des colonnes  $p_1, \dots, p_q$  de  $X_i$ , le modèle linéaire à effets mixtes s'écrit finalement :

►► **Modèle Structurel**, modèle pour la variable d'intérêt avec les effets aléatoires :

$$\forall i \leq N, \quad Y_i = X_i \beta + Z_i \eta_i + \varepsilon_i \quad (\text{MODLIN STR})$$

►► **Modèle Statistique**, modèle pour les effets aléatoires :

$$\forall i \leq N, \quad \eta_i \stackrel{iid}{\sim} \mathcal{N}(0, \Omega) \quad (\text{MODLIN STAT})$$

►► **Modèle d'Observation**, modèle pour les erreurs de mesures et le lien avec l'observation éventuellement :

$$\forall i \leq N, \quad \varepsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, \Sigma_i) \quad (\text{MODLIN OBS})$$

(MODLIN)

Nous pouvons étendre cette définition au modèle linéaire généralisé en considérant la valeur attendue de  $Y_i$ , à une transformation près,  $h(\mathbb{E}[Y_i])$ , comme observation.

### Modèles non-linéaires à effets mixtes

En réalité, les phénomènes biologiques sont rarement décrivables à l'aide d'équations linéaires. Si nous prenons par exemple l'exemple de dynamique anticorps après un second boost de vaccin, comme simulé en figure 9.1 ou observé en figure 9.7, cette dynamique ne peut être décrite, même à une transformation près, par un modèle linéaire. L'utilisation de modèle multipente pourrait permettre la modélisation de ce phénomène, mais d'autres enjeux comme l'estimation des points de cassure se posent alors [32]. Nous aspirons plutôt ici à mettre en évidence la limitation des modèles linéaires simples face à la complexité des dynamiques biologiques observées. Nous voulons pouvoir décrire des dynamiques biologiques non forcément linéaires, tout en incluant des connaissances à priori sur les mécanismes biologiques sous-jacents ainsi que des aspects mécanistes, afin de développer une approche plus adaptée à notre contexte spécifique.

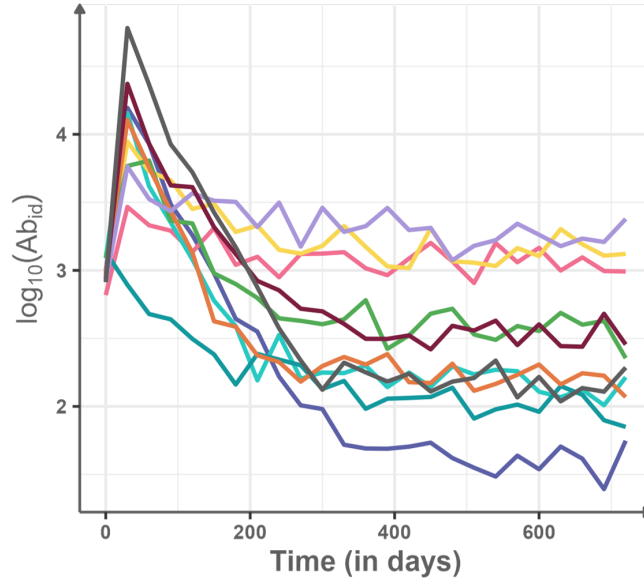


FIGURE 9.1 – Trajectoires individuelles simulées de la quantité d’anticorps observée dans la réponse immunitaire humorale à un premier boost de vaccin contre Ebola.

Maintenant, nous considérons une variable d’intérêt  $Y_i = (Y_{ij})_{j \leq n_i}$  observée aux temps  $t_{ij}$  pour l’individu  $i \leq N$ . Nous supposons que ces observations proviennent d’une dynamique individuelle  $y_i = f(\cdot, \psi_i)$  dépendante du temps et de paramètres  $\psi_i = (\psi_{il})_{l \leq m}$  du modèle. La fonction  $f$  est quelconque ici. Nos observations s’écrivent alors

$$\forall i \leq N, j \leq n_i, \quad Y_{ij} = f(t_{ij}, \psi_i) + \varepsilon_{ij}$$

avec  $\varepsilon_i = (\varepsilon_{ij})_{j \leq n_i} \stackrel{iid}{\sim} \mathcal{N}(0, \Sigma_i)$ .

Dans le cadre des modèles non-linéaires à effets mixtes, les paramètres individuels sont supposés distribués normalement  $\psi_i \stackrel{iid}{\sim} \mathcal{N}(\psi_{pop}, \Omega)$  autour d’un paramètre de population  $\psi_{pop} \in \mathbb{R}^m$ . De plus, si l’individu  $i$  est caractérisé par les covariables  $X_i = (X_{i1}, \dots, X_{in})$ , nous pouvons ajouter dans la définition des paramètres individuels un effet de covariables, de sorte que

$$\forall i \leq N, \quad \psi_i = \psi_{pop} + X_i \beta + \eta_i$$

avec  $\beta = (\beta_1, \dots, \beta_m)$ , où  $\beta_l = (\beta_{l1}, \dots, \beta_{ln})^T$  est l’effet des  $n$  covariables sur le paramètre  $l \leq m$ . Cette définition des paramètres  $\psi_i$  utilise donc un modèle linéaire à effets mixtes, mais il est possible d’utiliser un modèle linéaire généralisé à effets mixtes, avec une transformation  $h$ , de sorte que

$$\forall i \leq N, \quad h(\psi_i) = h(\psi_{pop}) + X_i \beta + \eta_i$$

Par exemple, une transformation logarithmique  $h = \log$  permet de contraindre les paramètres à être strictement positifs.

Nous pouvons finalement écrire le modèle non-linéaire à effets mixtes général comme

►► Modèle Structurel :

$$\forall i \leq N, j \leq n_i \quad \begin{cases} Y_{ij} &= f(t_{ij}, \psi_i) + \varepsilon_{ij} \\ h(\psi_i) &= h(\psi_{pop}) + X_i \beta + \eta_i \end{cases} \quad (\text{MOD STR})$$

►► Modèle Statistique :

$$\forall i \leq N, \quad \eta_i \stackrel{iid}{\sim} \mathcal{N}(0, \Omega) \quad (\text{MOD STAT})$$

►► Modèle d'Observation :

$$\forall i \leq N, \quad \varepsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, \Sigma_i) \quad (\text{MOD OBS})$$

(MOD)

où la valeur des  $(\psi_i)_{i \leq N}$  peut être estimée, si on prend par exemple par moindres carrés, comme

$$\forall i \leq N, \quad \hat{\psi}_i = \arg \min_{\psi \in \mathbb{R}^m} \sum_{j \leq n_i} (Y_{ij} - f(t_{ij}, \psi))^2$$

Il est important de noter qu'il est possible de ne définir qu'une partie seulement des paramètres comme individuels, et non forcément la totalité des paramètres.

### *Estimation des paramètres de populations d'un modèle non-linéaire à effet mixtes*

Comme expliqué précédemment, les modèles non-linéaires à effets mixtes offrent un cadre puissant pour modéliser des phénomènes complexes en tenant compte à la fois de l'hétérogénéité entre les individus et des relations non-linéaires au sein d'un modèle. Pour inférer les paramètres utilisés dans le modèle, l'estimation de maximum de vraisemblance peut être utilisée. Celle-ci consiste à trouver les valeurs des paramètres qui maximisent la vraisemblance, ou probabilité d'observer les données, définie par

$$\mathcal{L}_Y(\theta) = p(Y|\theta)$$

où  $Y$  correspond aux observations et  $\theta$  aux paramètres. Néanmoins, l'utilisation d'estimations par maximum de vraisemblance dans les modèles non-linéaires à effets mixtes peut présenter des limites importantes. Les modèles non-linéaires impliquent souvent des équations différentielles ordinaires et des interactions complexes entre les paramètres, ce qui peut rendre la fonction de vraisemblance difficile à calculer analytiquement. Les approximations numériques nécessaires pour résoudre ces équations et obtenir la vraisemblance peuvent introduire des incertitudes dans les estimations, notamment pour les queues de distribution, ce qui à terme risque de propager l'incertitude. De plus, le calcul de la vraisemblance peut être intensif en termes de temps et de ressources, rendant l'estimation coûteuse, particulièrement pour de grands ensembles de données [22].

C'est pour cela que d'autres approches ont été développées, comme la méthode NIMROD (Inférence par Approximation Normale dans les Modèles avec Effets Aléatoires basée sur les Équations Différentielles Ordinaires) [30]. NIMROD adopte une approche différente en utilisant une approximation normale de

la distribution postérieure des paramètres. Cette approximation, justifiée asymptotiquement par le théorème de Bernstein-Von Mises [36], implique de calculer le maximum de la distribution postérieure tout en approximant la matrice de variance-covariance à l'aide de l'inverse de la Hessienne du logarithme de la distribution postérieure. NIMROD offre un compromis entre la précision de l'estimation par maximum de vraisemblance et la rapidité de calcul de l'approximation normale, rendant cette méthode pertinente lorsque le calcul exact de la vraisemblance est intensif en termes de calcul ou que les approximations numériques introduisent des incertitudes.

Les méthodes bayésiennes implémentées dans des environnements tels que STAN, quant à elles, offrent une alternative pour l'estimation des paramètres dans les modèles non-linéaires à effets mixtes [6]. Elles permettent de prendre en compte des distributions a priori et d'obtenir des intervalles de confiance pour les paramètres, ce qui est utile notamment lorsque l'on s'intéresse aux queues de distribution. Elles reposent sur la représentation probabiliste de l'incertitude des paramètres et permettent l'incorporation d'informations a priori, ce qui peut stabiliser les estimations en cas de données limitées. En utilisant des méthodes de Monte Carlo Markov Chain (MCMC) [26, 15], la plateforme STAN échantillonne la distribution postérieure des paramètres, générant ainsi des échantillons représentatifs de l'ensemble des valeurs possibles. Les avantages incluent la possibilité d'estimer l'incertitude des paramètres via des intervalles de confiance, particulièrement pertinents pour des paramètres moins probables. Cependant, les méthodes bayésiennes peuvent nécessiter des temps de calcul importants et la spécification des distributions a priori peut influencer les résultats.

Finalement, une approche couramment utilisée pour l'estimation des paramètres dans les modèles non linéaires à effets mixtes est l'algorithme Stochastic Approximation Expectation-Maximization (SAEM), implémenté dans l'application Monolix [18, 23], application propriétaire accessible gratuitement à des fins académiques. L'algorithme permet l'estimation des paramètres en combinant l'algorithme d'espérance-maximisation [9] et de l'approximation stochastique. Tel qu'implémenté dans Monolix, l'algorithme comporte deux phases. Le but de la première phase est d'atteindre rapidement un voisinage de la solution en quelques itérations seulement. Une version de l'algorithme SAEM utilisant un recuit simulé, une méthode d'optimisation utilisée pour explorer l'espace des solutions d'un problème d'optimisation [17], accélère ce processus lorsque la valeur initiale est éloignée de la solution réelle. La seconde phase vise la convergence vers le maximum local trouvé, avec un comportement de plus en plus déterministe, semblable à un algorithme de gradient.

L'algorithme consiste essentiellement à construire, à l'itération  $k$ , à partir des observations  $Y = (Y_i)_{i \leq N}$ ,  $N$  chaînes de Markov  $(\psi_1^{(k)})$ , ...,  $(\psi_N^{(k)})$  qui convergent vers les distributions conditionnelles  $p(\psi_1|Y_1), \dots, p(\psi_N|Y_N)$ , en utilisant à chaque étape  $k \geq 1$  les données complètes  $(Y, \psi^{(k)})$  pour calculer un nouveau vecteur de paramètres  $\theta_k$ . L'algorithme SAEM a été démontré comme étant efficace pour un large éventail de modèles complexes, et il est particulièrement adapté pour les situations où les calculs analytiques exacts sont difficiles à obtenir.

### Modélisation de la réponse immunitaire humorale au boost vaccinal contre Ebola

Pour modéliser la réponse immunitaire humorale après le boost, nous considérons les dynamiques cellulaires décrites par le schéma présenté en figure 9.2, d'après le modèle de Pasin [27]. Les anticorps, notés  $Ab$  sont produits au taux  $\delta_{Ab}$  par des cellules sécrétrices d'anticorps caractérisées par leur durée de vie. Un premier type de ces cellules a une durée de vie courte, avec une demi-vie moyenne de 3 jours, et est noté  $S$ , tandis qu'un second type a une durée de vie plus longue, avec une demi-vie moyenne de 6 ans, et est noté  $L$ ; les demi-vies sont représentées par  $\frac{\ln(2)}{\delta_X}$  avec  $X \in \{S, L\}$ . Chacune de ces entités

a respectivement un taux de décroissance noté  $\delta_{Ab}$ ,  $\delta_S$  et  $\delta_L$ .

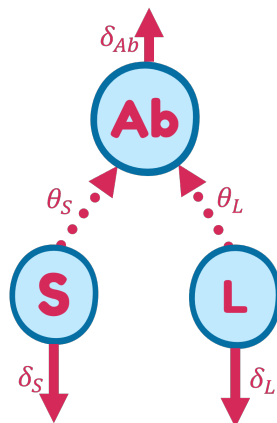


FIGURE 9.2 – Modèle de la réponse immunitaire humorale : les cellules sécrétrices d'anticorps Short-live (S) et Long-live (L) sécrètent respectivement aux taux  $\theta_S$  et  $\theta_L$  des anticorps (Ab). Ces entités biologiques se désagrègent respectivement aux taux  $\delta_S$ ,  $\delta_L$  et  $\delta_{Ab}$ .

Si nous considérons donc ces dynamiques, nous sommes capables d'écrire les quantités de chacune de ces entités,  $S(t)$ ,  $L(t)$ , et  $Ab(t)$ , au cours du temps à l'aide du système (9.1).

$$\left\{ \begin{array}{l} \frac{d}{dt} Ab(t) = \theta_S S(t) + \theta_L L(t) - \delta_{Ab} Ab(t) \\ \frac{d}{dt} S(t) = -\delta_S S(t) \\ \frac{d}{dt} L(t) = -\delta_L L(t) \\ (S, L, Ab)(t=0) = (S_0, L_0, Ab_0) \end{array} \right. \quad (9.1)$$

Le temps initial  $t = 0$  correspond ici à 7 jours après la deuxième dose de vaccin, ce dernier correspond théoriquement à l'état d'équilibre atteint par les cellules  $S$  et  $L$ . En notant  $\varphi_S = \theta_S S_0$  et  $\varphi_L = \theta_L L_0$ , nous pouvons finalement décrire la dynamique des anticorps à l'aide des paramètres  $\psi = (\varphi_S, \varphi_L, \delta_S, \delta_L, \delta_{Ab})$  et de l'unique équation suivante :

$$\left\{ \begin{array}{l} \frac{d}{dt} Ab(t) = \varphi_S e^{-\delta_S t} + \varphi_L e^{-\delta_L t} - \delta_{Ab} Ab(t) \\ Ab(t=0) = Ab_0 \end{array} \right. \quad (9.2)$$

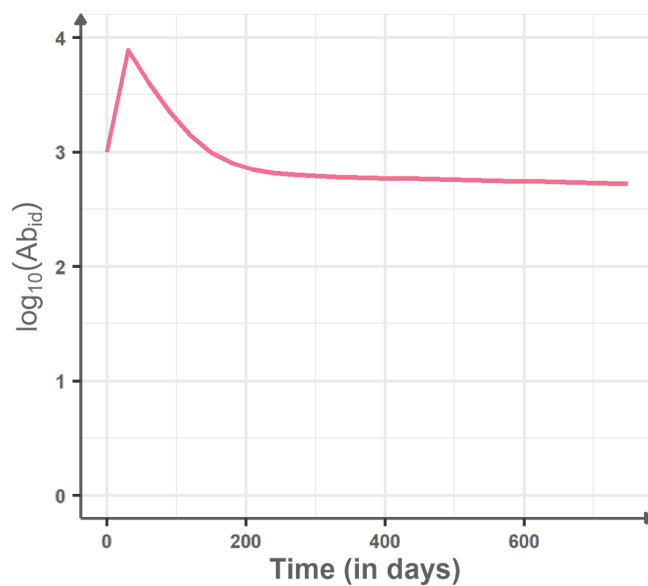


FIGURE 9.3 – Trajectoire moyenne simulée de la dynamique des anticorps dans la réponse immunitaire humorale à un premier boost de vaccin contre Ebola.

La figure 9.3 représente la trajectoire de ces anticorps pour les paramètres moyens donnés en annexe C.3. Si maintenant nous considérons la réponse immunitaire humorale de  $N$  patients et non plus d'un seul, la dynamique tracée sur la figure 9.3 peut ne pas correspondre à celle de chaque individu. En réalité, nous aurions tendance à avoir des trajectoires individuelles variées, comme en figure 9.4 sous l'effet des covariables de chaque individu, mais aussi des effets aléatoires propres à chaque individu.

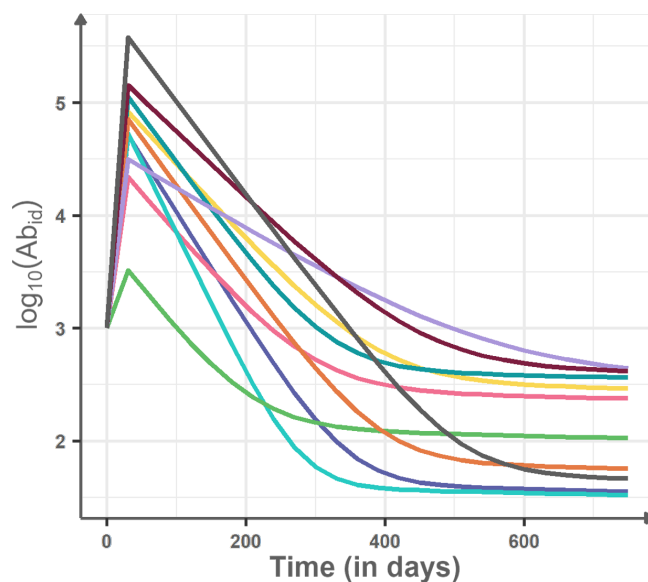


FIGURE 9.4 – Trajectoires individuelles simulées de la dynamique des anticorps dans la réponse immunitaire humorale à un premier boost de vaccin contre Ebola.

Pour l'exemple des dynamiques anticorps après un boost de vaccin contre Ebola, et d'après le travail de C. Pasin [27] et M. Alexandre [1], la variabilité individuelle dépend notamment des paramètres  $\varphi_S$ ,

$\varphi_L$  et  $\delta_{Ab}$  qui sont respectivement partiellement expliqués par l'âge, l'origine et le sexe des participants. Chacun de ces paramètres est décrit comme :

$$\forall i \leq N, \begin{cases} \log(\varphi_{Si}) &= \log(\varphi_{S_{pop}}) + \log(FC\Delta AGE = +1an) \times (AGE_i - \overline{AGE}) + \eta_i^S \\ \log(\varphi_{Li}) &= \log(\varphi_{L_{pop}}) + \beta_{RACE} \times \mathbb{1}_{RACE_i=European} + \eta_i^L \\ \log(\delta_{Abi}) &= \log(\delta_{Ab_{pop}}) + \beta_{SEX} \times \mathbb{1}_{SEX_i=Male} + \eta_i^{Ab} \end{cases}$$

où  $AGE_i$ ,  $RACE_i$  et  $SEX_i$  correspondent respectivement à l'âge, l'origine et le sexe du participant  $i$ . Les coefficients  $\beta_{RACE}$  et  $\beta_{SEX}$  correspondent à l'effet des covariables  $RACE$  et  $SEX$  sur, respectivement,  $\varphi_L$  et  $\delta_{Ab}$ ;  $FC\Delta AGE = +1an$  correspond, lui, aux facteurs multiplicatifs pour une différence d'âge de 1 an, et donc  $\log(FC\Delta AGE = +1an)$  représente l'effet de la différence d'âge à l'âge moyen sur le paramètre  $\varphi_S$ .

Finalement, le modèle utilisé pour la dynamique anticorps après un premier boost de vaccin contre Ebola est

►► Modèle Structurel :

$$\forall i \leq N, j \leq n_i, \begin{cases} \frac{d}{dt} Ab_i(t_{ij}) &= \varphi_{Si} e^{-\delta_S t_{ij}} + \varphi_{Li} e^{-\delta_L t_{ij}} - \delta_{Abi} Ab_i(t_{ij}) \\ Ab_i(t_{i0} = 0) &= Ab_{0i} \end{cases}$$

$$\begin{cases} \log(\varphi_{Si}) &= \log(\varphi_{S_{pop}}) + \log(FC\Delta AGE = +1an) \times (AGE_i - \overline{AGE}) + \eta_i^S \\ \log(\varphi_{Li}) &= \log(\varphi_{L_{pop}}) + \beta_{RACE} \times \mathbb{1}_{RACE_i=European} + \eta_i^L \\ \log(\delta_{Abi}) &= \log(\delta_{Ab_{pop}}) + \beta_{SEX} \times \mathbb{1}_{SEX_i=Male} + \eta_i^{Ab} \end{cases} \quad (\text{MOD AB STR})$$

►► Modèle Statistique :

$$\forall i \leq N, \begin{cases} \eta_i^S &\stackrel{iid}{\sim} \mathcal{N}(0, \omega_S^2) \\ \eta_i^L &\stackrel{iid}{\sim} \mathcal{N}(0, \omega_L^2) \\ \eta_i^{Ab} &\stackrel{iid}{\sim} \mathcal{N}(0, \omega_{Ab}^2) \end{cases} \quad (\text{MOD AB STAT})$$

►► Modèle d'Observation :

$$\forall i \leq N, j \leq n_i, \begin{cases} Y_{ij} &= \log_{10}(Ab_i(t_{ij}) + \varepsilon_{ij}) \\ \varepsilon_i &\stackrel{iid}{\sim} \mathcal{N}(0, \Sigma = \sigma_{Ab}^2 I_{n_i}) \end{cases} \quad (\text{MOD AB OBS})$$

(MOD AB)

Ici le modèle d'observation permet également de lier nos observations  $Y_{ij}$  à la variable d'intérêt  $Ab_i(t_{ij})$  selon la transformation  $h = \log_{10}$ . Les paramètres finaux du modèle sont  $\delta_S$ ,  $\delta_L$ ,  $\varphi_{S_{pop}}$ ,  $\varphi_{L_{pop}}$ ,  $\delta_{Ab_{pop}}$ ,  $\omega_S$ ,  $\omega_L$ ,  $\omega_{Ab}$  et  $\sigma_{Ab}^2$ .

### 9.1.2. Algorithmes de construction de modèle

Dans le domaine de la modélisation statistique, la sélection de modèle est une étape essentielle visant à identifier le modèle le plus approprié pour représenter au mieux les données et à réaliser des prédictions

précises. Pour faciliter ce processus, plusieurs algorithmes ont été développés. La sélection de modèle et l'utilisation d'algorithmes appropriés, tels que SCM (Stepwise Covariate Modeling) [34], COSSAC (COnditional Sampling use for Stepwise Approach based on Correlation tests) [2] et SAMBA (Stochastic Approximation for Model Building Algorithm) [31] sont d'une grande pertinence dans le cadre de la médecine personnalisée. La médecine personnalisée vise à adapter les traitements médicaux en fonction des caractéristiques individuelles de chaque patient, telles que son profil génétique, son état de santé, son mode de vie et ses réactions spécifiques aux médicaments. Dans ce contexte, la modélisation statistique et la sélection de covariables jouent un rôle crucial pour développer des modèles prédictifs permettant de personnaliser les traitements efficacement en fonction des caractéristiques, i.e. des covariables, de chaque individu.

### sélection de modèles

Un modèle complexe est capable de capturer de manière précise et détaillée les relations dans les données, y compris les interactions et les fluctuations mineures. Cependant, une complexité excessive, i.e. un trop grand nombre de paramètres, peut conduire à un surajustement *-overfitting-* du modèle aux données d'entraînement, ce qui signifie que le modèle est trop spécifique et ne généralise pas bien aux nouvelles données non vues. Des problèmes d'identifiabilité des paramètres peut également subvenir lorsque des modèles trop complexes sont employés. Le surajustement se traduit par une très bonne performance sur les données d'entraînement mais une performance prédictive mauvaise. D'un autre côté, un modèle simple a moins de paramètres et est donc plus général. Il est moins susceptible de surajuster les données d'entraînement, ce qui signifie qu'il peut mieux généraliser aux nouvelles données. Cependant, un modèle trop simple peut sous-représenter la complexité réelle des données et conduire à un sous-ajustement *-underfitting-* où le modèle ne capture pas bien les relations dans les données.

Le but de la sélection de modèles est donc de trouver le juste équilibre entre ces deux extrêmes. En utilisant des critères tels que l'AIC, le BIC et le BICc, défini ci-dessous, il est alors possible de pénaliser la complexité des modèles en ajoutant des termes qui dépendent du nombre de paramètres à estimer. Ces critères encouragent à choisir des modèles qui sont suffisamment complexes pour capturer les motifs importants dans les données tout en évitant une complexité excessive qui pourrait conduire à un surajustement.

Le BICc (BIC corrigé) est une variante du BIC. Il est spécifiquement adapté aux modèles à effets mixtes, en prenant à la fois en compte le nombre de paramètres de population avec le nombre total d'observations, mais aussi les paramètres individuels avec le nombre d'individus spécifiquement.

#### Définition 9.1.1 - BIC corrigé (BICc)

Soit un modèle  $\mathcal{M}$ , avec  $k$  paramètres estimés par  $\hat{\theta}$ . On suppose que  $\theta = (\theta_R, \theta_F)$  où  $\theta_R$  correspond aux paramètres associés aux effets aléatoires ou aux effets fixes des covariables impliqués dans un modèle aléatoire pour les paramètres individuels, et  $\theta_F$  correspond aux autres paramètres du modèle. Alors, en notant  $N$  le nombre d'individus et  $n_{tot}$  le nombre total d'observations, le critère BICc est défini par

$$BICc(\mathcal{M}) = \dim(\theta_R) \ln(N) + \dim(\theta_F) \ln(n_{tot}) - 2\mathcal{LL}(\hat{\theta}) \quad (\text{BICc})$$

pour les données  $(Y_i)_{i \leq N}$

Nous nous concentrons ici sur la sélection des covariables  $X_{i1}, \dots, X_{in}$ , pour chaque paramètre  $\psi_I$ ,



$l \leq m$ , utilisées pour définir le modèle linéaire généralisé

$$h_l(\psi_{il}) = h_l(\psi_{pop\ l}) + X_i\beta_{il} + \eta_{il}$$

SCM, *Stepwise Covariate Modeling* [16, 34]

Le Stepwise Covariate Modeling (SCM [16], et SCM+ [34]) est une méthode statistique utilisée dans le domaine de la pharmacocinétique et de la pharmacodynamique. Le principe de l'algorithme est décrit en annexe C.1.

L'algorithme SCM, bien qu'applicable à un grand nombre de modèle et accessible, présente également certaines limites. Le SCM peut être sensible à l'ordre dans lequel les covariables sont testées, ce qui signifie que différentes séquences de tests peuvent conduire à des modèles finaux différents. Les tests de rapport de vraisemblance utilisés dans le SCM reposent sur des hypothèses spécifiques concernant la distribution des résidus et peuvent ne pas toujours être appropriés pour toutes les situations.

COSSAC, *Conditional Sampling use for Stepwise Approach Based on Correlation tests* [2]

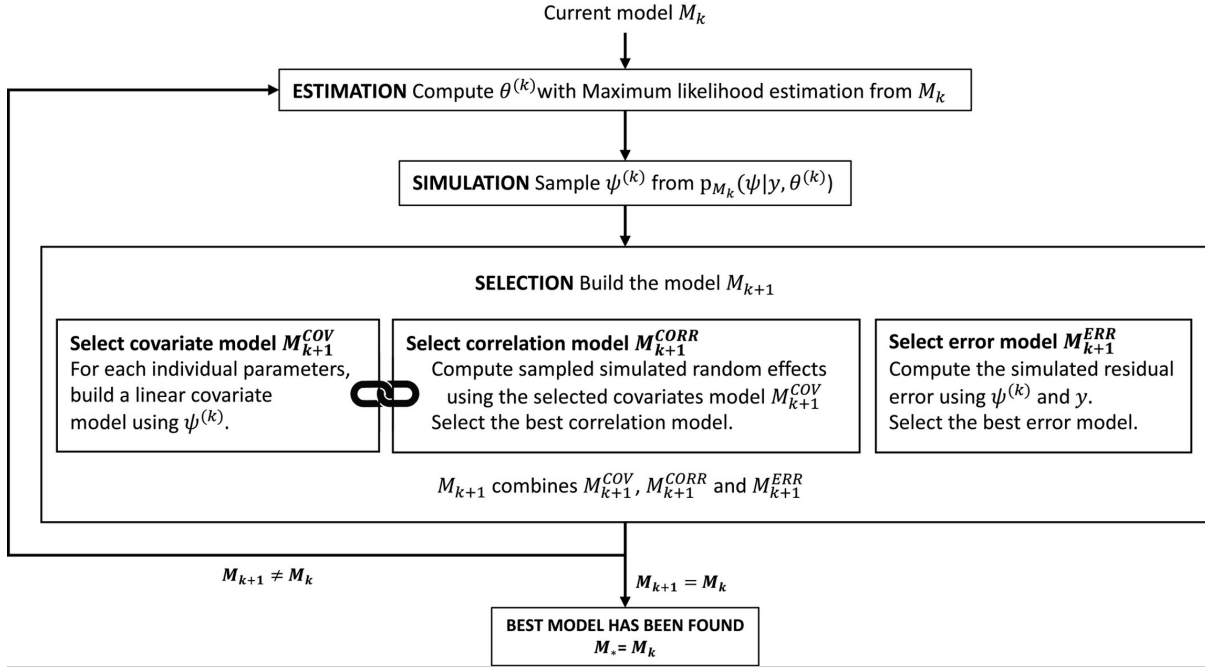
L'algorithme COnditional Sampling use for Stepwise Approach based on Correlation tests (COSSAC) [2] a été développé pour automatiser le processus de construction de modèle de covariables en identifiant les relations significatives entre les covariables et les paramètres du modèle. L'algorithme COSSAC vise à réduire le nombre d'itérations nécessaires pour construire un modèle covariable, tout en maintenant des performances similaires à d'autres méthodes traditionnelles, telles que le SCM précédemment introduit, en termes de qualité de modèle. Le principe de l'algorithme est décrit en annexe C.1.

Néanmoins, le choix du seuil de significativité pour les tests de corrélation peut influencer les covariables sélectionnées. Un seuil trop élevé peut entraîner la non-détection de relations potentiellement pertinentes, tandis qu'un seuil trop bas peut entraîner la sélection de covariables non significatives. De plus, l'algorithme COSSAC dépend de la qualité du modèle de base utilisé pour générer les échantillons des distributions conditionnelles. Un modèle de base inapproprié peut affecter la performance de l'algorithme COSSAC. Enfin, l'algorithme COSSAC est basé sur des corrélations linéaires entre covariables et paramètres, ce qui peut sous-estimer les relations non linéaires ou complexes.

SAMBA, *Stochastic Approximation for Model Building Algorithm* [31]

L'algorithme Stochastic Approximation for Model Building Algorithm (SAMBA) [31] a été développé pour accélérer et optimiser le processus de construction de modèles dans le domaine des modèles mixtes non-linéaires. L'objectif principal de SAMBA est de permettre de trouver rapidement des modèles de haute qualité en termes d'ajustement aux données et de parcimonie, même lorsque le nombre de paramètres ou de covariables disponibles est important.

Le principe de l'algorithme SAMBA, décrit en figure 9.5, repose sur l'idée d'apprendre des informations à partir d'un modèle "mauvais" tout en ajustant les données avec un modèle incomplet. Contrairement aux méthodes traditionnelles qui nécessitent de réestimer le modèle complet à chaque itération, SAMBA construit un nouveau modèle à chaque étape en utilisant les résultats obtenus avec le modèle précédent.



CPT Pharmacom & Syst Pharma, Volume: 11, Issue: 2, Pages: 161-172, First published: 01 February 2022, DOI: (10.1002/psp4.12742)

FIGURE 9.5 – Vue d’ensemble de l’algorithme Stochastic Approximation for Model Building Algorithm (SAMBA), extrait de l’article [31]

Le processus SAMBA est itératif. L’algorithme SAMBA commence avec un modèle initial qui peut être vide  $\mathcal{M}_0$  -c’est à dire ne contenant aucun effet de covariable, et aucune corrélation. Ensuite à chaque itération  $k$ , l’algorithme construit un nouveau modèle  $\mathcal{M}_{k+1}$  à partir du précédent  $\mathcal{M}_k$ . Pour cela, il estime les paramètres  $\theta^{(k)}$  du modèle  $\mathcal{M}_k$  par maximum de vraisemblance puis il échantillonne  $\psi_i^{(k)}$  pour chaque individu selon la loi  $p(\cdot|Y, \theta^{(k)})$  obtenu par Monte-Carlo. A partir de ces paramètres, l’algorithme construit trois sous-modèles qui constitueront  $\mathcal{M}_{k+1}$  :

- **Modèle de covariable**  $\mathcal{M}_{k+1}^{COV}$  : à partir des paramètres  $\psi_i^{(k)} = (\psi_{i1}^{(k)}, \dots, \psi_{im}^{(k)})$ ,  $i \leq N$ , générés, et  $\psi_{pop}^{(k)} = (\psi_{pop\ l}^{(k)})_{l \leq m}$  estimés, des tests statistiques exclus les covariables ne jouant significativement aucun rôle pour un paramètre. Ensuite, un modèle de régression est construit pour le paramètre  $\psi_l$ ,  $l \leq m$

$$h_l(\psi_{li}^{(k)}) = h_l(\psi_{lpop}^{(k)}) + c_i \beta_l + \eta_{il}^{(k)}$$

où  $\eta_{il}^{(k)} \sim \mathcal{N}(0, \omega_l^{(k)2})$ , et  $c_i$  est tout ou une partie des covariables de l’individu  $i$ , sélectionnées par une procédure stepwise, ”stepAIC” [37] .

- **Modèle de corrélation**  $\mathcal{M}_{k+1}^{COR}$  est estimé à partir des résidus des modèles de paramètres individuels estimés  $\eta_i^{(k)} = (\eta_{il}^{(k)})_{l \leq m}$ .
- **Modèle d’erreur**  $\mathcal{M}_{k+1}^{ERR}$  est estimé à partir des résidus du modèle structurel

$$e_{ij} = Y_{ij} - f(t_{ij}; \psi_i^{(k)})$$

En utilisant les résultats du modèle précédent pour guider la construction du modèle suivant, SAMBA réduit considérablement le nombre d’itérations et de réestimations du modèle complet, ce qui entraîne un temps de calcul nettement inférieur par rapport à des méthodes plus exhaustives. De plus, des études de comparaison ont montré que l’algorithme SAMBA produit des modèles similaires à ceux obtenus avec des méthodes plus coûteuses en termes de temps de calcul, telles que le SCM, tout en offrant des avantages de vitesse de calcul.

L'algorithme SAMBA est donc une méthode efficace pour accélérer le processus de construction de modèles dans les modèles mixtes non linéaires. Il offre des avantages significatifs en termes de vitesse de calcul, tout en produisant des modèles de haute qualité. Notre but pour ce mémoire, est d'améliorer le processus de sélection de covariables en grande dimension dans l'algorithme SAMBA, dans le but de mieux analyser des données de grandes dimensions et mieux déterminer les prédictors en grande dimension. Pour cela, une approche par LASSO est envisagée.

---

## 9.2. Sélection du modèle de covariables par LASSO dans SAMBA

---

### 9.2.1. Least Absolute Shrinkage and Selection Operator (LASSO)

Le LASSO, Least Absolute Shrinkage and Selection Operator, est une technique de régression linéaire pénalisée utilisée en statistiques. Il a été proposé en 1996 par le statisticien Robert Tibshirani [35].

Le LASSO est conçu pour résoudre le problème de sélection de variables et pour effectuer une régularisation dans les modèles de régression linéaire lorsque le nombre de variables explicatives est potentiellement beaucoup plus grand que le nombre d'observations. Dans de tels cas, les modèles traditionnels de régression linéaire peuvent être instables et sur-ajuster les données, ce qui conduit à une mauvaise prédiction ensuite.

#### Modèle et estimation par LASSO

Soit  $(X_i, Y_i)_{i \leq N}$  nos données, tels que  $X_i = (X_{i1}, \dots, X_{in})^T \in \mathbb{R}^n$  et  $Y_i \in \mathbb{R}$ . Nous supposons de plus que les  $(Y_i)_{i \leq N}$  sont indépendants, et que  $(X_i)_{i \leq N}$  sont standardisés, i.e. centré et réduit.

##### Définition 9.2.1 -

Dans le cadre linéaire, nous cherchons  $(\alpha, \beta)$  tel que :

$$\forall i \leq N, \quad Y_i = \alpha + X_i \beta + \varepsilon_i$$

avec  $\varepsilon$  une erreur. Nous définissons l'estimateur LASSO comme :

$$(\hat{\alpha}, \hat{\beta}) = \arg \min_{(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}^n} \left\{ \sum_{i \leq N} (Y_i - \alpha - X_i \beta)^2 \right\} + \lambda \|\beta\|_1 \quad (9.3)$$

où  $\lambda > 0$  est un paramètre de pénalisation à déterminer.

Pour tout  $\lambda > 0$ , la solution pour le paramètre  $\alpha$  est  $\hat{\alpha} = \bar{Y} = \frac{1}{N} \sum_{i \leq N} Y_i$ .  $(Y_i)_{i \leq N}$  est donc supposée, sans perte de généralité, telle que  $\bar{Y} = 0$ . Nous supposons dans un premier temps que le paramètre  $\lambda$  est

fixé, et nous détaillerons en partie 9.2.1 un algorithme pour le déterminer.

**Remarque :** Par dualité du Lagrangien, le problème d'optimisation (9.3) est équivalent au problème d'optimisation sous contrainte :

$$(\hat{\alpha}, \hat{\beta}) = \arg \min_{(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}^n, \|\beta\|_1 \leq t} \left\{ \sum_{i \leq N} (Y_i - \alpha - X_i \beta)^2 \right\} \quad (9.4)$$

### Pénalisation par Cross-Validation

Le choix de la pénalité  $\lambda$  dans la définition 1 est crucial pour la modélisation et fortement dépendant des données. Si le paramètre  $\lambda$  est trop faible, la pénalité sera négligeable et l'estimateur LASSO se comportera presque comme une régression linéaire ordinaire. Si en revanche le paramètre  $\lambda$  est trop élevé, de nombreux coefficients peuvent être réduits à zéro, ce qui conduit à une sous-modélisation, i.e. trop peu de prédictors dans le modèle.

Une méthode pour déterminer cette pénalité est de considérer une liste  $(\lambda_l)_{l \leq L}$  de paramètre  $\lambda$ , et tester lequel permet d'obtenir une meilleur prédiction. Le critère utilisé pour déterminer la qualité en terme de prédiction d'un modèle est appelé l'erreur de prédiction.

#### Définition 9.2.2 -

Pour un modèle de régression, l'erreur de prédiction est défini comme :

$$PE = \mathbb{E} [(Y - \hat{Y})^2] \quad (\text{PE})$$

où  $Y$  est l'outcome et  $\hat{Y}$  la prédiction utilisant le modèle.

Nous définissons également l'erreur quadratique moyenne de prédiction par

$$MSE = \frac{1}{N} \sum_{i \leq N} (Y_i - \hat{Y}_i)^2 \quad (\text{MSE})$$

Pour chacun des paramètres  $\lambda_l$  possible, il faut alors estimer l'erreur de prédiction (PE) du modèle  $\mathcal{M}_l$  pour  $\lambda = \lambda_l$ . La valeur de  $\lambda_l$  obtenant la plus faible valeur est alors choisi comme paramètre de pénalisation. Un algorithme utilisé pour estimer efficacement l'erreur de prédiction des modèles  $(\mathcal{M}_l)_{l \leq L}$  est un algorithme de validation croisée  $K$ -folds, détaillée en annexe C.2.

### Sélection de variables par LASSO

Le LASSO introduit une pénalité  $L^1$ , avec  $\|x\|_1 = \sum_{i \leq n} x_i$ ,  $x \in \mathbb{R}^n$  dans la fonction de coût, ce qui entraîne une régularisation du modèle. Cette pénalité favorise la sparsité, c'est-à-dire qu'elle tend à mettre à zéro les coefficients associés aux variables les moins importantes, permettant ainsi d'obtenir un modèle plus simple et plus interprétable, ce qui présente un grand intérêt dans la sélection de variables donc.

Cette propriété de sparsité ammené de la pénalisation  $L^1$  vient de la forme géométrique de la région de contrainte qu'elle crée. Si on prends l'exemple en deux dimensions, avec donc un  $\beta = (\beta_1, \beta_2) \in \mathbb{R}^2$ , la pénalité  $\lambda \|\beta\|_1$  est donc égal à  $\lambda_1 |\beta_1| + \lambda_2 |\beta_2|$ . La zone de contrainte alors créée pour les valeurs correspond à un diamand. Si on prenait la pénalisation  $L^2$ , la zone de contrainte serait alors décrit par un disque. Comme représenté sur la figure 9.6, si maintenant nous traçons, représenté en rouge, la

valeur de la fonction coût pour différente valeur de  $\beta$ , nous obtenons une forme ellipsoïdale autour de l'estimateur des moindres carrés non pénalisé. Comme on peut le voir sur la figure, en fonction de l'axe des ellipsoïdes sont alors plus susceptibles de croiser la zone de contrainte en un coin que pour la pénalité  $L^2$ . C'est en ces coins que les coefficients sont égaux à zéro.

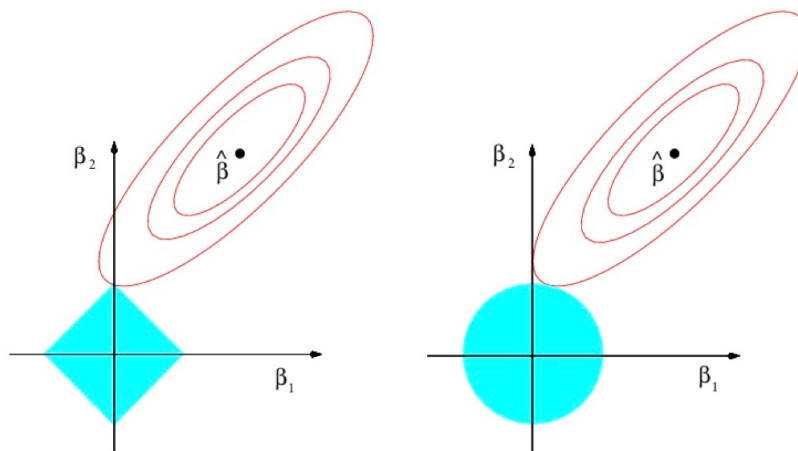


FIGURE 9.6 – Estimation de la fonction de coût (en rouge) et la pénalité (en bleu) dans le cas du LASSO (à gauche) et de la régression ridge, avec une pénalisation  $L^2$  (à droite).

La géométrie du LASSO induit donc une certaine sparsité dans les coefficients du modèle. Cette particularité du LASSO en fait un excellent choix pour la sélection de variables, puisqu'il tend à annuler certains coefficients et à conserver donc uniquement les coefficients les plus importants. En éliminant les coefficients associés aux variables moins importantes, le LASSO permet de réduire la dimensionnalité du problème également. Cela est particulièrement utile lorsque le nombre de variables est élevé par rapport à la taille de l'ensemble de données, ce qui pourrait entraîner un surajustement si toutes les variables étaient incluses.

Pour implémenter la méthode de sélection par LASSO, nous utilisons le package `glmnet` [13]. Ce package R réalisé par R. Tibshirani et al. propose une fonction `cv.glmnet` qui calcule le paramètre  $\lambda$  de pénalisation optimisant le critère MSE, à partir des données  $X$  et  $Y$ . Une seconde fonction utilisant `glmnet` utilise les données  $X$  et  $Y$ , ainsi que le paramètre  $\lambda$  choisi pour calculer l'estimateur par LASSO  $\hat{\beta}$  défini précédemment.

### 9.2.2. Implémentation du LASSO pour la sélection de variable dans SAMBA

Comme il a été présenté dans la partie 9.2.1, le LASSO est pertinent pour la sélection de variables en raison de sa capacité à fournir des estimateurs sparses, éliminant ainsi les variables moins importantes et favorisant l'inclusion des variables les plus pertinentes. Cela permet de simplifier le modèle, de réduire la dimensionnalité et de prévenir le surajustement en grande dimension. C'est pour ces raisons qu'une méthode par LASSO est essayée pour améliorer la procédure SAMBA. Nous allons donc décrire dans cette partie son application à un modèle non-linéaire à effets mixtes en vue de l'inclure dans l'algorithme SAMBA.

### Modèles pour les paramètres individuels

Nous reprenons les notations introduites en partie 9.1.1. Soit un modèle non-linéaire à effets mixtes quelconques (MOD). Soit  $N \in \mathbb{N}^*$  le nombre d'individus,  $n \in \mathbb{N}^*$  le nombre de covariables et  $m \in \mathbb{N}^*$  le nombre de paramètres du modèles, de sorte que :

- $\psi_i = (\psi_{i1}, \dots, \psi_{im})$  et  $\psi_{pop} = (\psi_{pop1}, \dots, \psi_{popm})$  sont des matrices lignes à  $m$  colonnes correspondant aux paramètres individuels et de population ;
- $X_i = (X_{i1}, \dots, X_{in})$  est une matrice ligne à  $n$  colonnes correspond aux covariables de l'individu  $i$  ;
- $\beta_l = (\beta_{l1}, \dots, \beta_{ln})^T \in \mathbb{R}^n$  correspond aux effets des covariables sur le paramètre  $l \leq m$ , et  $\beta = (\beta_1, \dots, \beta_m)$  est donc une matrice à  $n$  lignes et  $m$  colonnes ;
- $\eta_i = (\eta_{i1}, \dots, \eta_{im})$  est la matrice ligne des  $m$  effets aléatoires, de matrice de covariance  $\Omega \in \mathcal{M}_m(\mathbb{R})$ .

### Procédure de sélection par LASSO

Au cours de l'algorithme SAMBA, mentionné précédemment, des paramètres individuels sont généré à partir du modèle  $\mathcal{M}_k$  construit à l'itération précédente, et des paramètres  $\theta^{(k)}$  estimé pour  $\mathcal{M}_k$ . Ces paramètres  $\psi^{(k)}$  suivent donc le modèle

$$h_l(\psi_{il}^{(k)}) = h_l(\psi_{popl}^{(k)}) + X_i \beta_l + \eta_{il}^{(k)}$$

où  $\beta$  est donc à estimer. Pour cela, le modèle est écrit sous forme matriciel, en suivant la méthodologie proposé par [28], en posant

$$Y^{(k)} = \left( h_l(\psi_{il}^{(k)}) \right)_{i \leq N, l \leq m} = \mathbf{1}_N \left( h_l(\psi_{popl}^{(k)}) \right)_{l \leq m} + X\beta + \eta$$

où  $X = (X_1, \dots, X_i)^T \in \mathcal{M}_{Nn}(\mathbb{R})$ ,  $Y^{(k)} \in \mathcal{M}_{Nm}(\mathbb{R})$ ,  $\mathbf{1}_N = (1, \dots, 1)^T \in \mathbb{R}^N$ ,  $\eta = (\eta_1, \dots, \eta_N)^T \in \mathcal{M}_{Nm}(\mathbb{R})$ . Nous remarquons que les lignes de  $\eta_i$  de  $\eta$  sont indépendantes et identiquement distribuées selon  $\mathcal{N}(0, \Omega)$ . Dès lors, nous définissons le modèle de régression [38]

$$\mathcal{Y}^{(k)} = \mathcal{Y}_{pop}^{(k)} + \mathcal{X}\mathcal{B} + \mathcal{E}$$

avec

- $\mathcal{Y}^{(k)} = \text{vec}(Y^{(k)})$  et  $\mathcal{Y}_{pop}^{(k)} = \text{vec}(Y_{pop}^{(k)})$ , où  $X \in \mathcal{M}_{pq}(\mathbb{K}) \mapsto \text{vec}(X) \in \mathbb{K}^{p \times q}$  est l'opérateur vectorisant une matrice par ses colonne.
- $\mathcal{X} = (\Omega^{-1/2})^T \otimes X$  ;
- $\mathcal{B} = \text{vec}(\beta)$  ;
- $\mathcal{E} = \text{vec}(\eta\Omega^{-1/2})$

Finalement, on définit l'estimateur

$$\widehat{\mathcal{B}} = \arg \min_{\mathcal{B}} \left\{ \|\mathcal{Y}^{(k)} - \mathcal{Y}_{pop}^{(k)} - \mathcal{X}\mathcal{B}\|_2^2 + \lambda \|\mathcal{B}\|_1 \right\}$$

comme estimateur de  $(\beta_1, \dots, \beta_l)^T \in \mathbb{R}^{m \times n}$ . Dans le cadre de sélection de variable, les variables sélectionnées pour un certain paramètre  $l \leq m$  correspondent aux coefficient de  $\beta_l$  non mis à zéro.

### Algorithme de Stability Selection

Le Stability Selection Algorithm (l'algorithme de sélection par stabilité) [25] est une méthode développée pour rendre la sélection par LASSO plus stable et robuste. La sélection par LASSO est sujette

à des résultats instables car elle est basée sur des procédures stochastiques, telles que l'algorithme de validation croisée, qui peuvent donner des résultats différents d'un essai à l'autre.

L'idée centrale de l'algorithme de sélection par stabilité, écrit dans l'algorithme 8, est de répéter le processus de sélection par LASSO un grand nombre de fois, en utilisant différents échantillons de données (batches de données) tirés au hasard à chaque itération. À chaque itération, les paramètres du modèle sont estimés par LASSO en utilisant un échantillon aléatoire de taille  $\left\lfloor \frac{N}{2} \right\rfloor$  où  $N$  est la taille de l'ensemble de données. Cette approche de bootstrap permet de réduire l'impact des variations dues à la sélection d'un échantillon particulier.

Après avoir effectué ces nombreuses estimations, le nombre de fois où chaque covariable est sélectionnée dans le modèle final est comptabilisé. Une variable est conservée si sa fréquence de sélection, représentée par la proportion d'itérations où cette variable est non nulle, dépasse un seuil prédéfini  $t_{SS}$ . En général, un seuil de  $t_{SS} = 0.95$  est souvent utilisé

Cette approche de sélection par stabilité offre plusieurs avantages. Tout d'abord, elle permet de contrôler le taux de fausses découvertes, car seules les variables sélectionnées fréquemment sont incluses dans le modèle final. Deuxièmement, elle améliore la stabilité du processus de sélection, car les variations dues à l'échantillonnage sont lissées par la répétition des estimations.

---

**Algorithme 8 : Sélection par stabilité [25]**


---

**Data :** Soit un modèle  $\mathcal{M}$ , des données  $(X_i, Y_i)_{i \leq N}$ , un paramètre de pénalisation  $\lambda$  fixé. On définit aussi un paramètre  $n_{SS}$  et  $t_{SS}$  correspondant respectivement au nombre d'estimation et au seuil de sélection de l'algorithme.

**pour**  $k \leq n_{SS}$  **faire**

Les paramètres du modèle  $\mathcal{M}$  sont estimés par LASSO, pour le même paramètre  $\lambda$ , en utilisant les données  $(X_i, Y_i)_{i \in \mathcal{I}_k}$  où  $\mathcal{I}_k \subset \{1, \dots, N\}$  de cardinal  $\left\lfloor \frac{N}{2} \right\rfloor$  choisi aléatoirement (uniformément) ; un estimateur  $\beta_k$  est obtenu ;  
 Les paramètres  $\beta_l^k$  nulles sont comptabilisés ;

**fin**

Pour chaque paramètre  $\beta_l$ ,  $l \leq m$ , la covariable  $s \leq n$ , est sélectionné si la fréquence  $\nu$ , où le paramètre  $\beta_{ls}$  est non nulle, au cours des itérations précédentes est supérieur à  $t_{SS}$ .

---

### 9.2.3. Décroissance du critère de sélection au cours des itération

Dans l'algorithme SAMBA initial, la sélection de covariables par stepAIC assure la décroissance du critère utilisé pour la construction du modèle. De fait, au cours des itérations de l'algorithme, les modèles  $(\mathcal{M}_k)$  construit ont un critère AIC, BIC ou BICc selon le choix de l'utilisateur qui diminue au cours des itérations. En ayant recours à une méthode par LASSO, la décroissance du critère est perdue, et les modèles peuvent au cours des itérations aller vers un modèles de critère plus élevés et donc moins pertinents.

Dans le but d'éviter ce phénomène, nous ajoutons en plus du LASSO un test sur le critère à chaque itération selon le principe suivant :

- Soit le modèle  $\mathcal{M}_k$  retenu à la fin de l'itération  $k$ , on construit à partir de ce dernier un modèle de covariable  $\mathcal{M}_{cons}$ .
- Si le critère de ce dernier n'est pas meilleur que le précédent, alors le modèle  $\mathcal{M}_{k+1}$  est choisi égal

à  $\mathcal{M}_k$  et une nouvelle itération est lancée.

- Si le critère est meilleur que le précédent, alors le modèle construit  $\mathcal{M}_{cons}$  est choisi comme  $\mathcal{M}_{k+1}$ .

En faisant de la sorte, si un modèle n'améliore pas le critère, l'algorithme est relancé avec le modèle précédent, et non stoppé, permettant éventuellement d'en trouver un meilleur grâce à la stochasticité permise dans l'algorithme de recherche. En revanche si au bout de  $s = 3$  recherches, l'algorithme ne parvient pas à construire un modèle améliorant le critère, alors l'algorithme s'arrête, rendant le dernier modèle retenu comme modèle final.

---

### 9.3. Application à la dynamique humorale post-boost

---

Dans le but d'appliquer l'algorithme de sélection de modèle modifié, nous voulons évaluer les performances sur des simulations semblables au procédé d'étude de l'essai clinique Prevac-UP contre Ebola [3]. L'évaluation des performances des méthodes se traduisent par l'analyse du taux ou nombre de faux positifs, et faux négatifs.

#### 9.3.1. Cadre d'application

L'essai clinique Prevac-UP [3] a pour but de tester l'efficacité de plusieurs stratégies vaccinales. Au cours de celui-ci, des adultes (1400) et des enfants (1401) se sont vus administrer un vaccin contre Ebola, puis on été suivis dans les mois suivants. Pour chaque individu, des mesures d'anticorps, représentés en figure 9.7 ont été faites le jour de la première dose, puis 7, 14, 28, 56 et 63 jours après, puis 3, 6 et 12 mois après. La seconde dose a lieu 56 jours après la première.



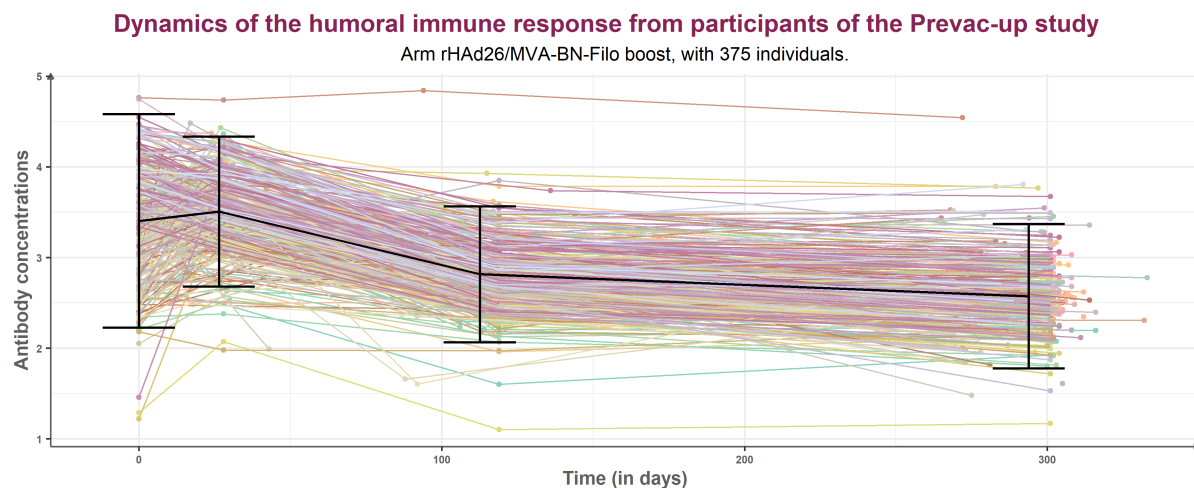


FIGURE 9.7 – Trajectoires individuelles de la réponse immunitaire humorale 7 jours après un premier boost des participants à l’essai clinique PREVAC-Up[3]; avec la trajectoire moyenne et les intervalles de confiance à 95% en chaque jour de visite théorique d’observation. La stratégie vaccinale est rHAd26/MVA-BN-Filo .

Au cours de l’essai Prevac-Up, une sous-étude immunologique a été menée sur 95 adultes. Au cours de celle-ci, différentes stratégies vaccinales ont été essayées puis des mesures d’expressions génétiques ont été collectées 56 jours après la première dose (i.e. le jour de la seconde dose) puis 63 jours, 3, 6 et 12 mois après la première dose.

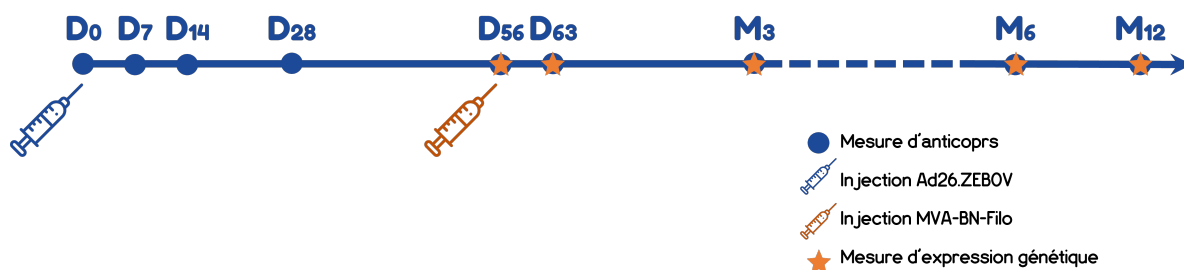


FIGURE 9.8 – Design de l’essai clinique PREVAC-Up, avec l’exemple de la stratégie vaccinale avec comme prime Ad26.ZEBOV et comme boost MVA-BN-Filo.

L’une des stratégies vaccinales mise en place est un vaccin contre Ebola basé sur Ad26.ZEBOV/MVA-BN-Filo, il s’agit d’un schéma à deux doses comprenant Ad26.ZEBOV comme première dose et MVA-BN-Filo comme deuxième dose. La dynamique post boost de ce vaccin a notamment été estimée et évaluée par C. Pasin [27], pour l’essai clinique EBOVAC [12]. Un modèle a alors été proposé pour modéliser la dynamique de la réponse immunitaire humorale, et il s’agit de celui présenté dans le modèle (MOD AB STR). C’est à partir de ce modèle et des valeurs estimées pour ce modèle par C. Pasin et M. Alexandre lors de travaux antérieurs que nous créons des données de simulations.

Pour ce faire, nous utilisons le logiciel Simulx [24] qui permet de générer facilement des dynamiques individuelles au sein d’une population, selon des covariables individuelles, mais aussi selon une certaine variabilité intraindividuelle. Nous avons généré 100 jeux de données, de 100 individus, dont les covariables significatives suivent les distributions :

$$— AGE_i \stackrel{iid}{\sim} \mathcal{U}([20; 50]);$$

- $SEX_i \stackrel{iid}{\sim} \mathcal{B}(0.5)$ , où 0 correspond à une femme et 1 à un homme ;
- $RACE_i \stackrel{iid}{\sim} \mathcal{B}(0.5)$ , où 0 représente un participant d'origine africaine et 1 un participant d'origine européenne ;

A partir de ces covariables, et en utilisant le modèle (MOD AB), avec les valeurs des paramètres estimées lors de travaux antérieurs sur l'essai clinique EBOVAC [12] et regroupées dans le tableau 1 en annexe C.3, nous générons des paramètres individuels pour chacun des 100 individus de chaque jeu de données. Nous ajoutons ensuite des covariables non pertinentes, avec ou sans corrélations, selon les formules et la méthode décrites en annexe C.4.

Maintenant que les covariables et paramètres de chaque individu sont générées, les observations bruitées des dynamiques individuelles obtenues sont simulées ; les temps d'observations sont 7, 14, 56, 63 jours après 7 jours post-boost, puis 6 et 12 mois après. Un exemple de dynamiques est tracé en figure 9.1.

### 9.3.2. Application à des simulations

Les méthodes décrites précédemment ont donc été appliquées aux simulations produites, pour 100 et 200 covariables d'abord, avec 100 individus.

#### Pertinence de l'algorithme de sélection par stabilité

Afin d'évaluer la pertinence de l'utilisation de l'algorithme de stabilité sélection décrit en partie 9.2.2, nous générons pour chaque individus  $i \leq N$  des covariables suivant les distributions suivantes :

- $AGE_i \stackrel{iid}{\sim} \mathcal{U}([20; 50])$  ;
- $SEX_i \stackrel{iid}{\sim} \mathcal{B}(0.5)$ , où 0 correspond à une femme et 1 à un homme ; ;
- $RACE_i \stackrel{iid}{\sim} \mathcal{B}(0.5)$ , où 0 représente un participant d'origine africaine et 1 un participant d'origine européenne ;

Nous rajoutons également des covariables non significatives pour le modèles, nommés *WEIGHT*, *SIZE*, *INCOME*, *STUDY*, *TENSION*, *SPORT*, *SMOKER*, qui sont indépendants entre elles et des précédentes.

A partir des covariables générées, nous simulons les paramètres individuels, en utilisant le modèle décrit par les équations MOD AB. Les paramètres  $\delta_S$  et  $\delta_L$  sont fixés pour tout les individus. Les valeurs utilisées pour les différentes paramètres individuels, de population, et d'erreur sont ceux résumés en tableau 1.

Nous avons donc simulé un jeu de données contenant des paramètres individuels  $(\psi_i) = (\varphi_{S\ i}, \varphi_{L\ i}, \delta_{AB\ i})_{i \leq N}$  et leurs covariables  $X_i = (AGE_i, SEX_i, RACE_i, \dots)$ . A ce jeu de données, nous appliquons l'algorithme de stabilité sélection 8, pour un paramètre  $\lambda$  de pénalisation estimée en amont par validation croisée 9. Nous obtenons les fréquences d'échantillonnage résumé en tableau 9.1. Comme nous pouvons le voir, peu de variables non significatives sont rarement sélectionnées ( $< 0.05$ ) mais en revanche, avec un seuil  $t_{SS} = 0.95$ , seule celle justement significative sont sélectionnées à l'issue des  $n_{SS} = 1000$  itérations ici. Bien que les variables non significatives ait une chance non négligeable d'être sélectionnées avec le LASSO, l'algorithme de sélection par stabilité permet de ne garder que les plus pertinentes.

|               | AGE            | INC.   | SIZE   | TEN.   | WEI.   | RACE          | SEX           | SMO.   | SPO.   | STU.   |
|---------------|----------------|--------|--------|--------|--------|---------------|---------------|--------|--------|--------|
| $\varphi_S$   | <u>0.9972</u>  | 0.0592 | 0.2163 | 0.3617 | 0.0119 | 0.0484        | 0.1051        | 0.1472 | 0.3699 | 0.0443 |
| $\varphi_L$   | 0.0887         | 0.1261 | 0.0776 | 0.0252 | 0.0146 | <u>1.0000</u> | 0.0870        | 0.1077 | 0.1435 | 0.2630 |
| $\delta_{Ab}$ | <u>0.04625</u> | 0.5088 | 0.3571 | 0.2070 | 0.1331 | 0.2980        | <u>0.9992</u> | 0.2541 | 0.5643 | 0.2313 |

TABLE 9.1 – Résultat de l'algorithme de stabilité sélection, avec  $n_{SS} = 1000$ , sur un jeu de données simulés. Les fréquences de sélection inférieure à 0.05 sont indiqués en clair ■, et celle supérieure à 0.95 sont indiqués en foncé ■. Les fréquences soulignés sont celles des bonnes associations variables-paramètres

*Application de la méthode par lasso avec algorithme de sélection par stabilité dans SAMBA*

La figure 9.9 présente les résultats pour les scénarios avec et sans corrélation pour 100 covariables, et 100 individus. Dans les scénarios avec et sans corrélations, le nombre de faux positifs diminue par la méthode par lasso. Le pourcentage de sélection des covariables non significatives passe de 9% initialement, avec corrélation, et 10%, sans corrélation, à 4% sans corrélation et 3% avec corrélation. Le pourcentage de sélection des covariables significatives, lui, est diminué de 5 à 6% en moyenne dans le scénario sans corrélation et de 6 à 12% dans le cas avec corrélation. La figure 20 présente les résultats en précisant les paramètres liés en annexe C.5. Le nombre de modèles construits avec aucun faux négatifs, c'est-à-dire aucune covariable significative manquante ou mal associée, augmente dans le cas sans corrélation lorsque la méthode par lasso est utilisée, passant de 78% à 81%. Le résultat avec corrélation diminue lui, de 78% à 72%.

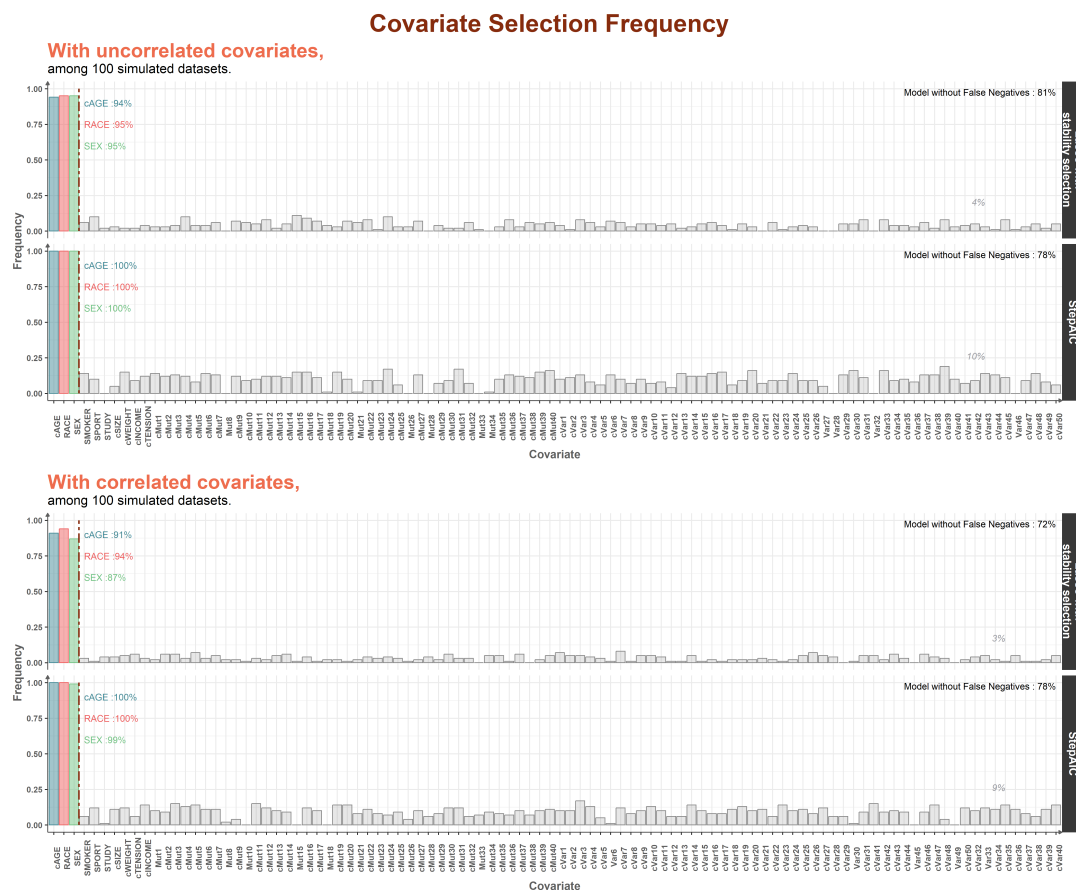


FIGURE 9.9 – Fréquence de sélection des différentes covariables, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations, pour 100 covariables, 100 individus.

Si maintenant nous augmentons le nombre de covariables à 200, avec toujours 100 individus, la figure 9.10 présente les résultats de sélection de modèles. De nouveau, le pourcentage de faux positifs diminue avec la méthode avec lasso, passant de 12% avec et sans corrélation en moyenne à 1% dans les deux cas. Néanmoins, la sélection de variables significatives diminue également, perdant jusqu'à 13% dans le scénario sans corrélation, et de 11% en moyenne avec corrélation. Le nombre de modèles finalement construits avec aucun faux négatif augmente dans le cas sans corrélation de 17% et diminue avec corrélation de 3%.

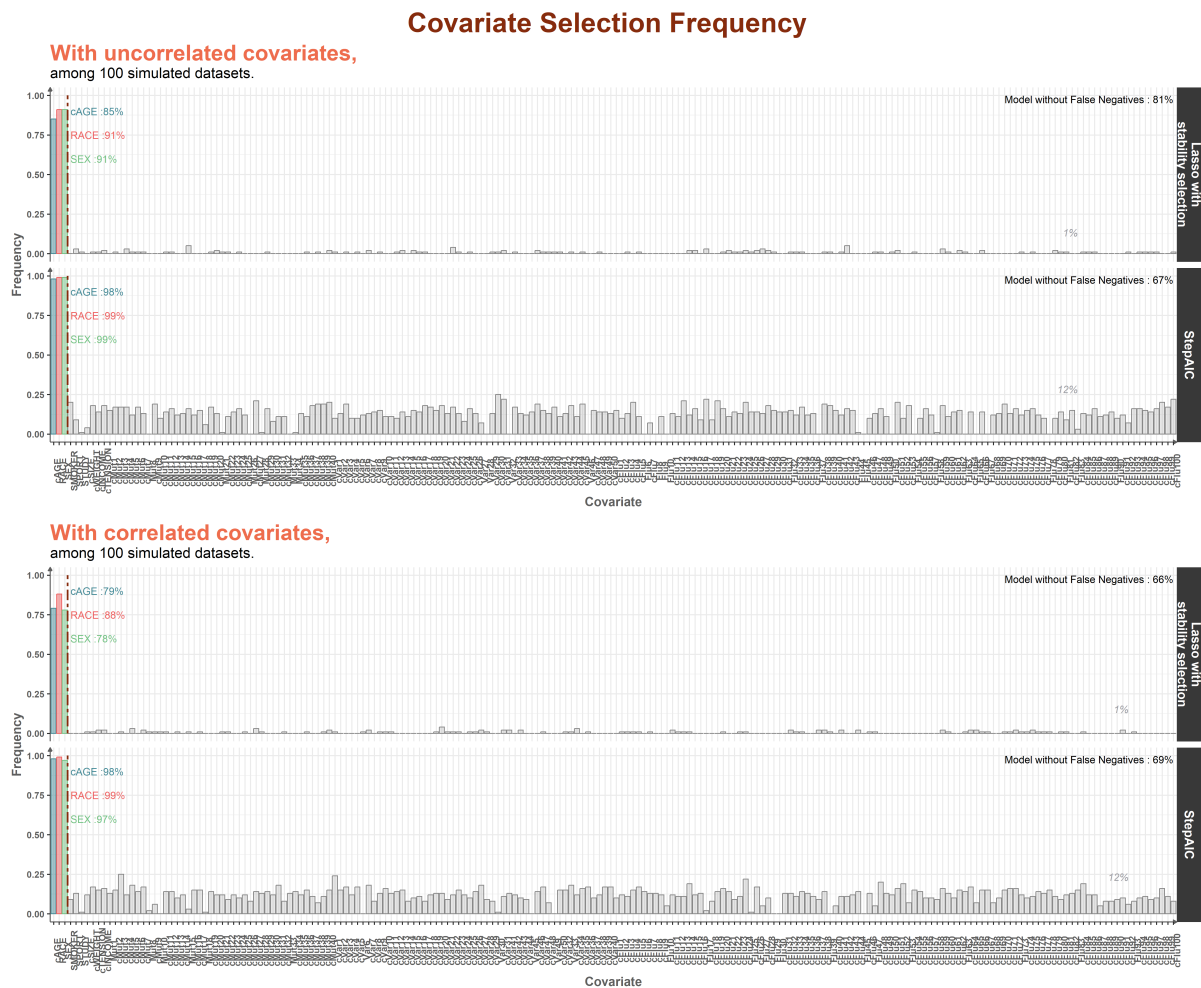


FIGURE 9.10 – Fréquence de sélection des différentes covariables, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations, pour 200 covariables, 100 individus.

### Application avec critère d'arrêt sur la décroissance du BICc

Le critère d'arrêt décrit en partie 9.2.3 a été implémenté dans SAMBA avec la méthode par lasso. Pour 100 covariables, les résultats en figure 9.11 montrent des résultats similaires avec la méthode sans le critère d'arrêt en terme d'amélioration sur les faux positifs. Le pourcentage de sélection des covariables non significatives reste diminué de 5-6% par rapport à la méthode initiale, mais cette amélioration n'est pas aussi bonne que sans ce critère. La sélection des variables significatives, elle, est toujours diminuée par rapport à la méthode initiale, mais relativement semblable aux résultats sans le critère d'arrêt. Les pourcentages de sélection changent de 1 ou 2% seulement par rapport à l'algorithme sans critère d'arrêt. Enfin, le pourcentage de modèles construits sans aucun faux négatif est toujours amélioré avec la nouvelle méthode, mais seulement de 1%, sans corrélation ; celui-ci est cependant diminuée, de 4%, avec corrélation.

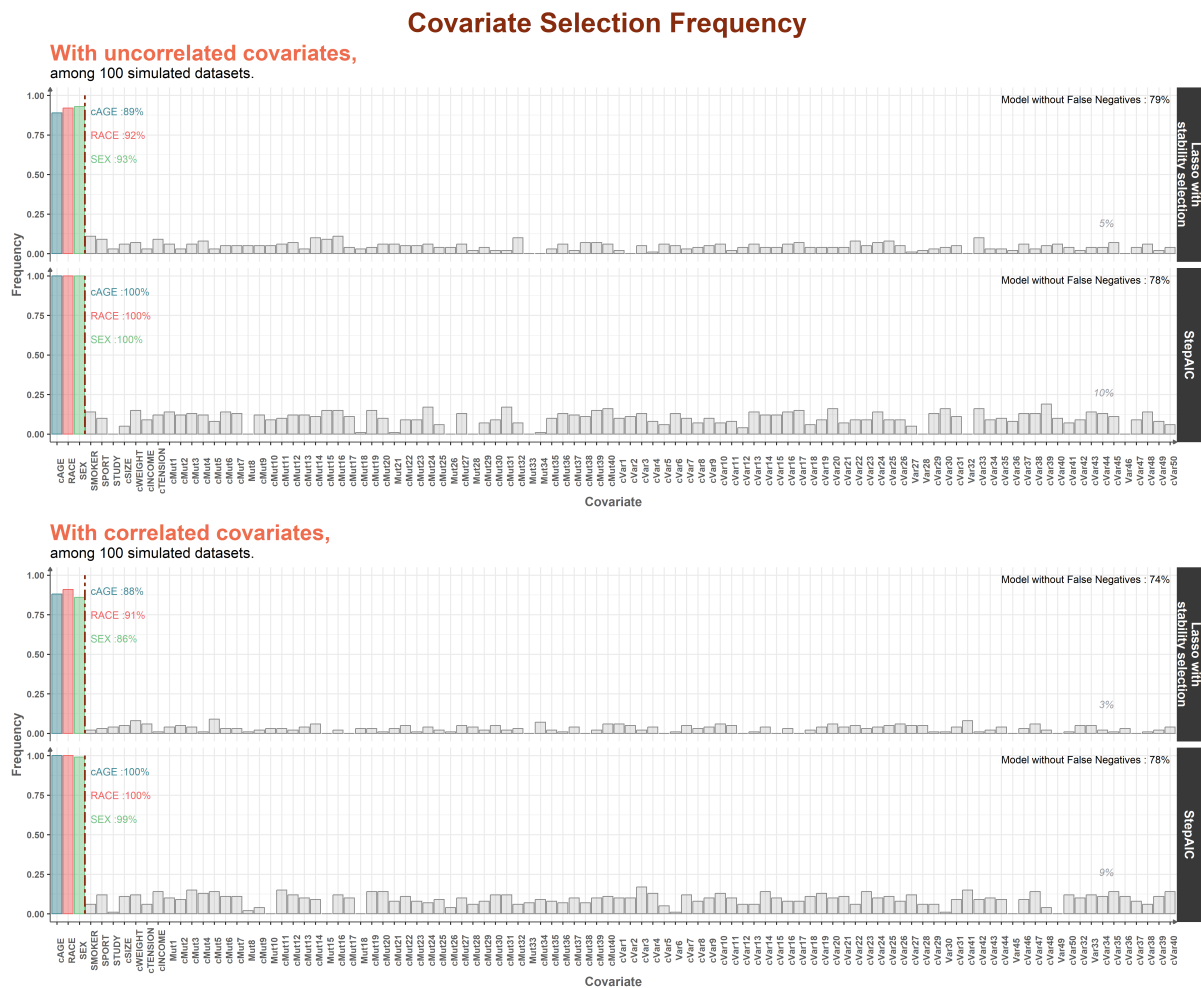


FIGURE 9.11 – Fréquence de sélection des différentes covariables, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations, pour 100 covariables, 100 individus ; le test d'arrêt sur la décroissance du critère, ici le (BICc), est appliqué.

Si le nombre de covariables est de 200, les résultats présentés en figure 9.12 montrent de meilleur résultat que précédemment. Le pourcentage de sélection de covariables non significatives passe de 12% avec et sans corrélation, pour la méthode initiale, à 1% comme précédemment. Le pourcentage de sélection des variables significatives est moins diminué que sans le critère d'arrêt. Maintenant, pour le scénario sans corrélation, le pourcentage de sélection de ces variables est diminué de 3 à 8%, contre 13% précédemment, et pour le scénario avec corrélation, le pourcentage est diminué de la même manière que précédemment, de 16% en moyenne. Dans la même idée, le pourcentage de modèles sans aucun faux négatif est augmenté 17% toujours, et est diminué de 2%, contre 3% précédemment, dans le cas avec corrélation.

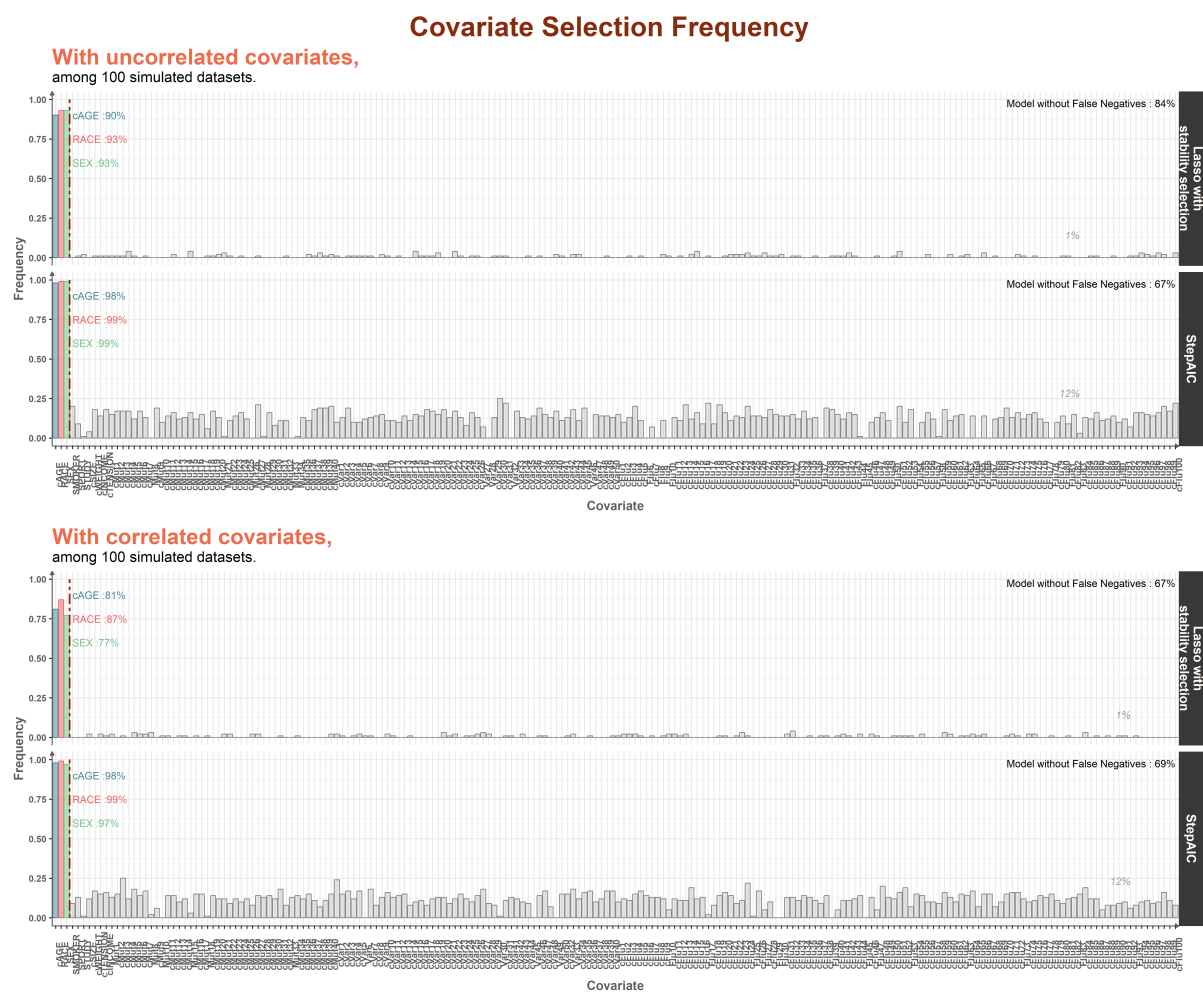


FIGURE 9.12 – Fréquence de sélection des différentes covariables, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations pour 200 covariables, 100 individus ; le test d'arrêt sur la décroissance du critère, ici le (BICc), est appliqué.

## False Discovery Rate

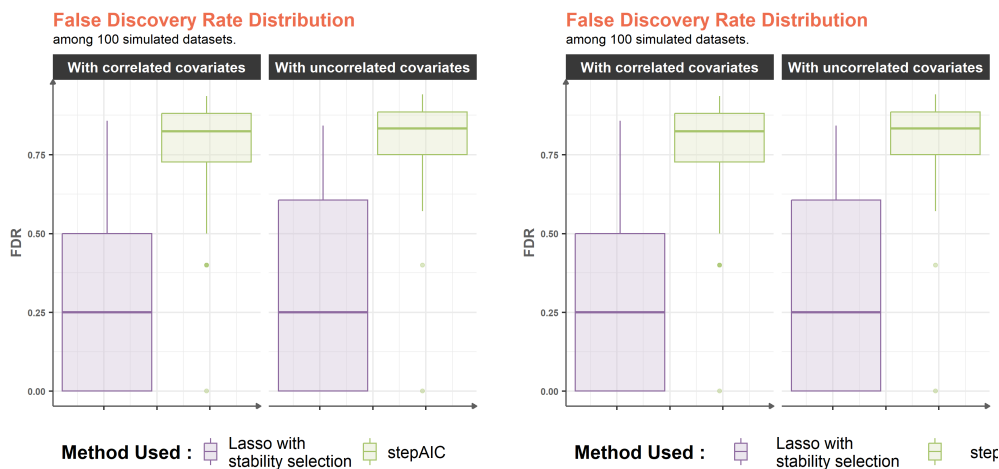


FIGURE 9.13 – Répartition du False Discovery Rate selon la méthode et la présence ou non de corrélations. Pour 100 covariables, 100 individus, à gauche et 200 covariables, 100 individus, à droite.

Lorsqu'on examine le nombre de faux positifs entre les deux méthodes, dont la répartition est présentée en figure 9.13, on observe que la méthode initiale, lorsqu'elle est utilisée avec 100 covariables, affiche un taux de faux positifs (FDR) de 80.6%, tandis qu'avec 100 covariables corrélées, ce taux diminue légèrement à 79.5%. En augmentant le nombre de covariables à 200, le FDR reste inchangé, s'élevant à 80.6% pour les covariables non corrélées et à 79.5% pour les covariables corrélées. En revanche, la nouvelle méthode présente une nette amélioration en termes de FDR. Avec 100 covariables, le FDR chute à 31.2%, et il est encore plus bas, à 28.2%, lorsque les covariables sont corrélées. Cette réduction des faux positifs souligne l'efficacité de la nouvelle méthode, surtout lorsqu'il est essentiel de minimiser le nombre de faux positifs, même avec une augmentation du nombre de covariables.

En comparant les performances de la méthode initiale et de la nouvelle méthode, la méthode initiale présente des taux de faux positifs moyens élevés, avoisinant 80%, quel que soit le nombre de covariables ou leur corrélation, suggérant un taux élevé de faux positifs. En contraste, la nouvelle méthode affiche des taux moyens significativement plus bas, environ 31% avec 100 covariables et 28% avec 200 covariables, indiquant une réduction des faux positifs. De plus, la borne supérieure à 90% du taux de faux positifs pour la nouvelle méthode reste en dessous de 80%, démontrant sa robustesse face à des scénarios exigeants. En conclusion, la nouvelle méthode se révèle être préférable pour minimiser les faux positifs et garantir des résultats plus fiables lors de l'analyse des données.



## Temps d'exécution

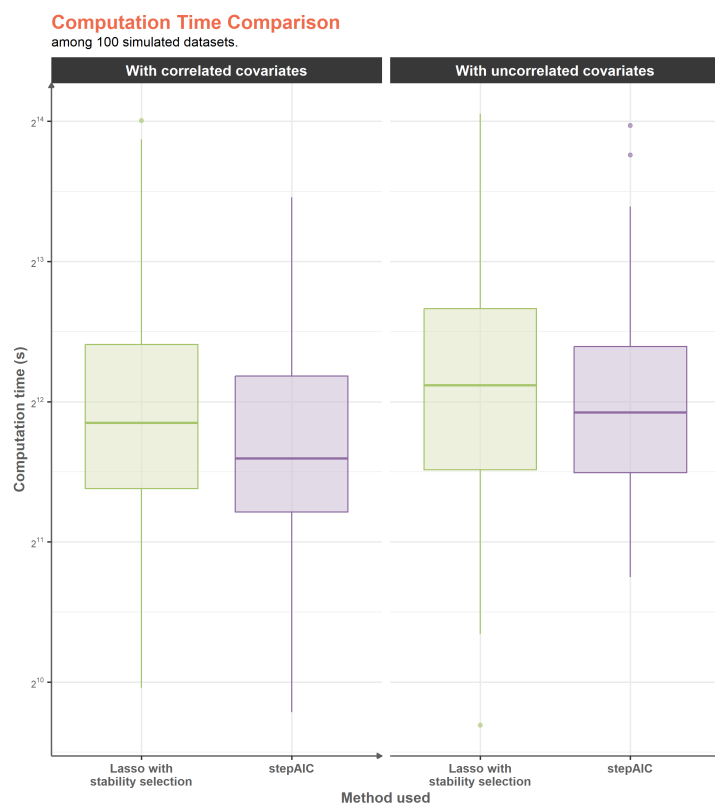


FIGURE 9.14 – Temps d'exécution de l'algorithme de construction du modèle final selon la méthode de sélection utilisée, pour 200 covariables et 100 individus.

En comparant les performances de deux algorithmes en termes de temps d'exécution sur la figure 9.14, il est clair que la méthode initiale de sélection de covariables est plus rapide que la nouvelle méthode, avec un modèle construit en moyenne en 3094 secondes lorsque des corrélations sont présentes, et en 3881 secondes en l'absence de corrélations. Cependant, le temps d'exécution n'augmente que de 14% en présence de corrélations et de 19% en l'absence de corrélations, en apportant une amélioration sur le nombre de faux positifs, le False Discovery Rate final obtenu, et le nombre de modèles sans aucun faux négatif. Ainsi, bien que la nouvelle méthode nécessite un temps d'exécution légèrement plus long, elle peut être préférable dans des contextes où le contrôle du nombre de faux positifs est plus important, même si cela se traduit par une augmentation de temps d'exécution de l'ordre de 15% en moyenne.

9.3.3. *Prevac-UP*

La sous-étude immunologie de l'essai clinique Prevac-UP [3] regroupe 95 participants, dont l'expression de 28 593 gènes a été récupéré en plus des mesures d'anticorps habituels, selon le schéma 9.8 présenté précédemment. Les trajectoires individuelles de la réponse immunitaire humorale de ces 95 individus seuls est représenté en figure 9.15.

| Parameter                                | Value   |         |
|------------------------------------------|---------|---------|
| Fixed Effects                            |         |         |
| delta_S_pop                              | 0.23    |         |
| delta_L_pop                              | 0.00032 |         |
| phi_S_pop                                | 1.31    |         |
| phi_L_pop                                | 1.34    |         |
| delta_AB_pop                             | 3.24    |         |
| Standard Deviation of the Random Effects |         |         |
|                                          | Value   | C.V.(%) |
| omega_phi_S                              | 0.27    | 27.19   |
| omega_phi_L                              | 0.28    | 28.4    |
| omega_delta_AB                           | 0.39    | 39.99   |
| Error Model Parameters                   |         |         |
| a                                        | 0.48    |         |

TABLE 9.2 – Valeur des différents paramètres du modèle (MOD AB STR) estimés par le logiciel Monolix.

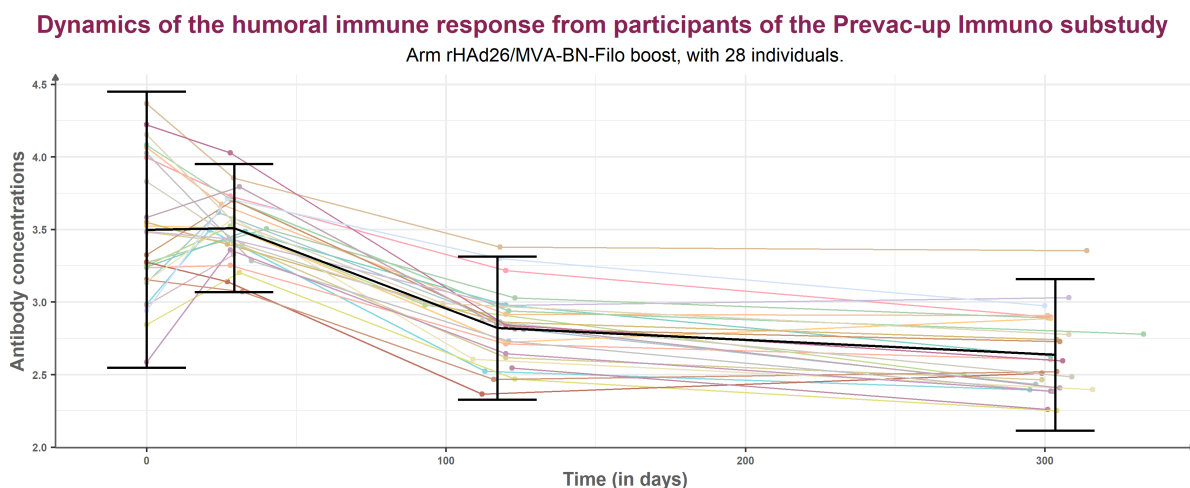


FIGURE 9.15 – Trajectoires individuelles de la réponse immunitaire humorale 7 jours après un premier boost des participants à la sous-étude immunologique de l’essai clinique PREVAC-Up[3] ; avec la trajectoire moyenne et les intervalles de confiance à 95% en chaque jour de visite théorique d’observation. La stratégie vaccinale est rHAd26/MVA-BN-Filo .

L’estimation des différents paramètres a été réalisé à l’aide du logiciel Monolix introduit précédemment. Les résultats sont présentés dans le tableau 9.2. Les valeurs de  $\delta_S$  et  $\delta_L$  ont été fixé par souci d’identifiabilité aux valeurs estimées dans les précédents travaux évoqués [1, 27]. L’ordre de grandeur des différents estimés est cohérent avec les travaux précédents également. Les trajectoires individuelles résultantes sont données en annexe C.6.

Nous souhaitons déterminer parmi les gènes mesurés à D63 les prédicteurs de la réponse immunitaire humorale à 7 jours post-boost. Nous avons donc d’abord tenté d’appliquer la méthode présentée sur le jeu de données résultant, et sur le jeu restreint aux 200 gènes ayant le plus de variabilité. Ces deux applications n’ont donné aucun résultat, que ce soit avec la méthode initiale ou la méthode par LASSO. L’idée a donc ensuite été de réduire d’abord la dimension du jeu de données initial. Pour faire cela, nous réduisons la dimension en utilisant une méthode de clusterisation `ClustOfVar` [8] reposant sur la

création d'un arbre hiérarchisant les groupes de variables corrélées dans le but de réduire la redondance d'information et de résumer chaque groupe en une variable numérique synthétique.

La méthodologie proposée dans cet article vise à résoudre les problèmes de réduction de dimension et de sélection de variables. La clusterisation de variables, nommée CoV, regroupe les variables hautement corrélées en clusters, chaque cluster étant résumé par une variable synthétique numérique. Cela permet de réduire la redondance d'informations et de conserver toutes les variables dans des clusters, simplifiant ainsi l'analyse. Ensuite, le nombre optimal de clusters est déterminé en fonction des performances de prédiction d'un classifieur basé sur des forêts aléatoires. Les variables synthétiques sont ensuite utilisées pour résumer l'information et comme jeux de données pour l'algorithme de construction de modèle SAMBA. L'utilisation de cette méthode demande beaucoup de ressources informatiques est en train d'être réalisé pour ce stage, dans l'objectif de l'appliquer avant la fin du stage.

---

## Liste des références

---

ALEXANDRE, M., PRAGUE, M., MCLEAN, C., BOCKSTAL, V., DOUGUIH, M. et THIÉBAUT, R. (2023). Evaluatoin and prediction of the long-term humoral immune response induved by the two-dose heterologous Ad26.ZEBOC,MVA-BN-Filo vaccine regimen against ebola. draft manuscript, pre-submitted.

AYRAL, G., SI ABDALLAH, J.-F., MAGNARD, C. et CHAUVIN, J. (2021). A novel method based on unbiased correlations tests for covariate selection in nonlinear mixed effects models : The cossac approach. *CPT : Pharmacometrics & Systems Pharmacology*, 10(4) :318–329.

BADIO, M., LHOMME, E., KIEH, M., BEAVOGUI, A., KENNEDY, S., DOUMBIA, S., LEIGH, B., SOW, S., DIALLO, A., FUSCO, D., KIRCHOFF, M., TERMOTE, M., VATRINET, R., WENTWORTH, D., ESPEROU, H., LANE, H., PIERSON, J., WATSON-JONES, D., ROY, C. et YAZDANPANA, Y. (2021a). Partnership for research on ebola vaccination (prevac) : protocol of a randomized, double-blind, placebo-controlled phase 2 clinical trial evaluating three vaccine strategies against ebola in healthy volunteers in four west african countries. *Trials*, 22.

BADIO, M., LHOMME, E., KIEH, M., BEAVOGUI, A., KENNEDY, S., DOUMBIA, S., LEIGH, B., SOW, S., DIALLO, A., FUSCO, D., KIRCHOFF, M., TERMOTE, M., VATRINET, R., WENTWORTH, D., ESPEROU, H., LANE, H., PIERSON, J., WATSON-JONES, D., ROY, C. et YAZDANPANA, Y. (2021b). Partnership for research on ebola vaccination (prevac) : protocol of a randomized, double-blind, placebo-controlled phase 2 clinical trial evaluating three vaccine strategies against ebola in healthy volunteers in four west african countries. *Trials*, 22.

BOËLLE, P.-Y. (2017). L'épidémie en guinée, sierra leone et libéria, 2014-2015. *Actualité et Dossier de Santé Publique*, 98.

CARPENTER, B., GELMAN, A., HOFFMAN, M. D., LEE, D., GOODRICH, B., BETANCOURT, M., BRUBAKER, M., GUO, J., LI, P. et RIDDELL, A. (2017). Stan : A probabilistic programming language. *Journal of statistical software*, 76(1).

CENTRE INRIA DE L'UNIVERSITÉ DE BORDEAUX, UNIVERSITÉ DE BORDEAUX, INSERM (n.d.). Sism team. <https://www.inria.fr/fr/sism>. Accessed on October, 2023.

CHAVENT, M., GENUER, R. et SARACCO, J. (2018). Combining clustering of variables and feature selection using random forests.

DEMPSTER, A. P., LAIRD, N. M. et RUBIN, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society : Series B*, 39 :1–38.

DUVIGNAUD, A., LHOMME, E., PISTONE, T., ONAISI, R., SITTA, R., JOURNOT, V., NGUYEN, D., PEIFFER-SMADJA, N., CRÉMER, A., BOUCHET, S., DARNAUD, T., POITRENAUD, D., PIROTH, L., BINQUET, C., MICHEL, J.-F., LEFÈVRE, B., LEBEAUX, D., LEBEL, J., DUPOUY, J., ROUSSILLON, C., GIMBERT, A., WITTKOP, L., THIÉBAUT, R., ORNE-GLIEMANN, J., JOSEPH, J.-P., RICHERT, L., ANGLARET, X., MALVY, D. et the COVERAGE study GROUP (2020). Home treatment of older people with symptomatic sars-cov-2 infection (covid-19) : A structured summary of a study protocol for a multi-arm multi-stage (mams) randomized trial to evaluate the efficacy and tolerability of several experimental treatments to reduce the risk of hospitalisation or death in outpatients aged 65 years or older (coverage trial). *Trials*, 21(1) :846.

EFRON, B. et TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap*. Numéro 57 de Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, Boca Raton, Florida, USA.

Eurosurveillance editorial TEAM, C. (2014). The innovative medicines initiative launches call on ebola and other filoviral haemorrhagic fevers. *Eurosurveillance*, 19(45).

FRIEDMAN, J., TIBSHIRANI, R. et HASTIE, T. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1) :1–22.

GOLDFELD, K. et WUJCIAK-JENS, J. (2020). simstudy : Illuminating research methods through data generation. *Journal of Open Source Software*, 5(54) :2763.

HASTINGS, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1) :97–109.

JONSSON, E. N. et KARLSSON, M. O. (1998). Automated Covariate Model Building Within NONMEM. *Pharmaceutical Research*, 15(9) :1463–1468.

KIRKPATRICK, S., GELATT, C. D. et VECCHI, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598) :671–680.

KUHN, E. et LAVIELLE, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational statistics & data analysis*, 49(4) :1020–1038.

LAIRD, N. M. et WARE, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38 4 :963–74.

LAVIELLE, M. (2014). *Mixed Effects Models for the Population Approach : Models, Tasks, Methods and Tools*. Chapman and Hall/CRC.

LAVIELLE, M. (2022). *Rsmix : R Speaks 'Monolix'*. R package version 5.0.1.

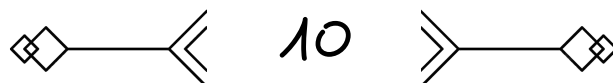
LEHMANN, E. L. et CASELLA, G. (1998). *Theory of Point Estimation*. Springer-Verlag, New York, NY, USA, second édition.

LIXOFT SAS, A SIMULATIONS PLUS COMPANY (2023a). Monolix. Version 2021R2.

LIXOFT SAS, A SIMULATIONS PLUS COMPANY (2023b). Simulx. Version 2021R2.

- MEINSHAUSEN, N. et BÜHLMANN, P. (2010). Stability selection. *Journal of the Royal Statistical Society, Series B*, 72 :417–473.
- METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H. et TELLER, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6) : 1087–1092.
- PASIN, C., BALELLI, I., VAN EFFELTERRE, T., BOCKSTAL, V., SOLFOROSI, L., PRAGUE, M., DOUGUIH, M. et THIÉBAUT, R. (2019). Dynamics of the Humoral Immune Response to a Prime-Boost Ebola Vaccine : Quantification and Sources of Variation. *Journal of Virology*, 93(18) :e00579–19.
- PERROT-DOCKÈS, M. (2019). *Méthodes régularisées pour l'analyse de données multivariées en grande dimension : théorie et applications*. Thèse de doctorat, Thèse de doctorat dirigée par Lévy-Leduc, Céline et Chiquet, Julien Mathématiques aux interfaces Université Paris-Saclay (ComUE) 2019.
- PERROT-DOCKÈS, M., LÉVY-LEDUC, C., SANSONNET, L. et CHIQUET, J. (2017). Variable selection in multivariate linear models with high-dimensional covariance matrix estimation.
- PRAGUE, M., COMMENGES, D., GUEDJ, J., DRYLEWICZ, J. et THIÉBAUT, R. (2013). Nimrod : A program for inference via a normal approximation of the posterior in models with random effects based on ordinary differential equations. *Computer Methods and Programs in Biomedicine*, 111(2) :447–458.
- PRAGUE, M. et LAVIELLE, M. (2022). Samba : A novel method for fast automatic model building in nonlinear mixed-effects models. *CPT : Pharmacometrics & Systems Pharmacology*, 11(2) :161–172.
- SEGALAS, C. (2019). *Inference for random changepoint models : application to pre-dementia cognitive decline*. Theses, Université de Bordeaux.
- SPEARMAN, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1) :72–101.
- SVENSSON, R. J. et JONSSON, E. N. (2022). Efficient and relevant stepwise covariate model building for pharmacometrics. *CPT : Pharmacometrics & Systems Pharmacology*, 11(9) :1210–1222.
- TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society (Series B)*, 58 :267–288.
- VAN DER VAART, A. (2000). *Asymptotic statistics*, volume 3. Cambridge university press.
- ZHANG, Z. (2016). Variable selection with stepwise and best subset approaches. *Annals of Translational Medicine*, 4(7).
- ZHU, W., LÉVY-LEDUC, C. et TERNÈS, N. (2021). A variable selection approach for highly correlated predictors in high-dimensional genomic data. *Bioinformatics*, 37(16) :2238–2244.

# Utilisation de modèle à variables latentes pour la construction de modèle mécaniste en grande dimension : application en immunologie pour le développement de vaccin



Auriane GABAUT  
Encadré par Mélanie PRAGUE et Cécile PROUST-LIMA  
Dès Octobre 2023

## Sommaire

---

|                                                                         |            |
|-------------------------------------------------------------------------|------------|
| <b>Introduction et Motivations</b>                                      | <b>211</b> |
| <b>10.1 Cadrage de la recherche</b>                                     | <b>212</b> |
| 10.1.1 Motivation                                                       | 212        |
| 10.1.2 Modèles non linéaire à effets mixtes                             | 213        |
| 10.1.3 Pertinence des modèles non-linéaires à effets mixtes en biologie | 215        |
| 10.1.4 Modèle à Variables Latentes                                      | 215        |
| 10.1.5 L'exemple de la réponse immunitaire humorale                     | 216        |
| <b>10.2 Analyse Méthodologique et Travail envisagé</b>                  | <b>218</b> |
| 10.2.1 Travail sur le modèle statistique                                | 218        |
| 10.2.2 Travail sur le modèle d'observation                              | 219        |
| 10.2.3 L'étude Prevac-UP, et autres applications                        | 220        |
| <b>10.3 Attendus</b>                                                    | <b>221</b> |
| 10.3.1 Calendrier des trois années de thèses                            | 221        |
| 10.3.2 Perspectives                                                     | 222        |

---

---

## Introduction et Motivations

---

La vaccination est l'une des mesures les plus efficaces pour prévenir les maladies infectieuses. Cependant, la réponse immunitaire à la vaccination varie grandement entre les individus, et les mécanismes sous-jacents ne sont pas entièrement compris. La modélisation mathématique offre un potentiel pour améliorer notre compréhension des réponses immunitaires induites par les vaccins et identifier les facteurs contribuant à la variabilité individuelle de ces réponses. Les modèles mécanistes sont des modèles mathématiques robustes permettant de prendre en compte cette variabilité individuelle tout en modélisant des phénomènes biologiques complexes. La dynamique des populations cellulaires en réponse au vaccin peut être analysée à l'aide de modèles mécanistiques basés sur des équations différentielles ordinaires. Cependant, il n'est pas toujours évident de déterminer quelles cellules spécifiques devraient être prises en compte dans ces modèles, et la mesure des populations cellulaires par cytométrie en flux représente une analyse coûteuse en ressources. En inférant les populations cellulaires à l'aide des expressions génétiques, il serait alors possible de mieux comprendre les dynamiques cellulaires propres à la réponse immunitaire humorale en incorporant ces données aux modèles mécanistes d'ores-et-déjà utilisés. Ce projet de thèse vise donc à développer une nouvelle approche pour construire des modèles mécaniste en grande dimension. L'intérêt et l'originalité du problème résident dans le fait qu'il cherche à combler le fossé entre la modélisation mécanistique et les données de grande dimension, un domaine qui n'a pas encore été suffisamment exploré dans la littérature scientifique. L'objectif de ce travail est d'utiliser un modèle à variables latentes qui établira des liens entre les marqueurs transcriptomiques observables et les variables ou processus latents représentant les mécanismes cellulaires et métaboliques impliqués dans la réponse immunitaire à la vaccination, qui sont, eux, non observés. L'hypothèse est qu'une approche mécanistique combinant des modèles basés sur des équations différentielles ordinaires avec des techniques d'apprentissage statistique permettra d'obtenir un modèle plus interprétable et robuste pour analyser les données transcriptomiques de grande dimension. Les enjeux de ce projet sont multiples. Premièrement, en évitant l'approche en deux étapes actuellement utilisée, qui consiste à sélectionner d'abord quelques marqueurs transcriptomiques avant de les inclure dans le modèle, nous espérons obtenir des associations entre les marqueurs plus pertinentes et moins biaisées ; de même, on s'attend à réduire la propagation de l'incertitude. Deuxièmement, cette méthodologie pourrait permettre de mieux analyser des données de grandes dimensions d'ores-et-déjà disponibles, comme dans la sous-étude immuno de Prevac-UP [12]. Cette thèse permettrait, en combinant des approches mécanistiques et statistiques, d'améliorer notre compréhension des réponses immunitaires aux vaccins. Les résultats attendus contribueront non seulement à l'avancement des connaissances dans le domaine de l'immunologie, mais pourraient également avoir des implications importantes pour le développement de vaccins plus efficaces, fournir de nouveaux outils de modélisation applicables pour d'autres domaines de recherche confrontés à des problèmes similaires de données de grande dimension et de modélisation mécanistique.

## 10.1. Cadrage de la recherche

### 10.1.1. Motivation

La méthode proposée dans ce projet de recherche revêt une grande pertinence dans le contexte actuel où des jeux de données transcriptomiques de grandes dimensions sont de plus en plus disponibles. Ces données, riches en informations, présentent cependant des défis considérables en termes d'analyse, nécessitant des outils robustes et novateurs pour une compréhension approfondie. Actuellement, la pratique courante consiste à sélectionner d'abord quelques marqueurs transcriptomiques avant de les intégrer dans un modèle mécanistique, ce qui peut entraîner une propagation d'incertitudes et une perte de précision dans l'analyse. En revanche, la méthode proposée dans cette thèse vise à aborder ces problèmes de front en combinant l'étape de sélection de marqueurs et la modélisation en une seule étape intégrée. Cela permettra d'identifier directement les associations pertinentes entre les marqueurs et les processus cellulaires sous-jacents, réduisant ainsi les biais et améliorant la précision des résultats. De plus, l'utilisation de cette approche intégrée est rendue d'autant plus pertinente par l'évolution récente des techniques expérimentales. Auparavant, obtenir des données de grandes dimensions était un défi technique majeur. Cependant, grâce au Coverage Immuno Project mené par l'équipe SISTM, il a été démontré que des mesures de l'expression de 20 000 gènes à chaque point temporel étaient possibles en utilisant des gouttes de sang prélevées par des tests de piqûres [5]. Cette avancée technologique ouvre de nouvelles perspectives pour la recherche en immunologie, mais elle nécessite également des méthodes d'analyse robustes pour exploiter pleinement ces données riches en informations.

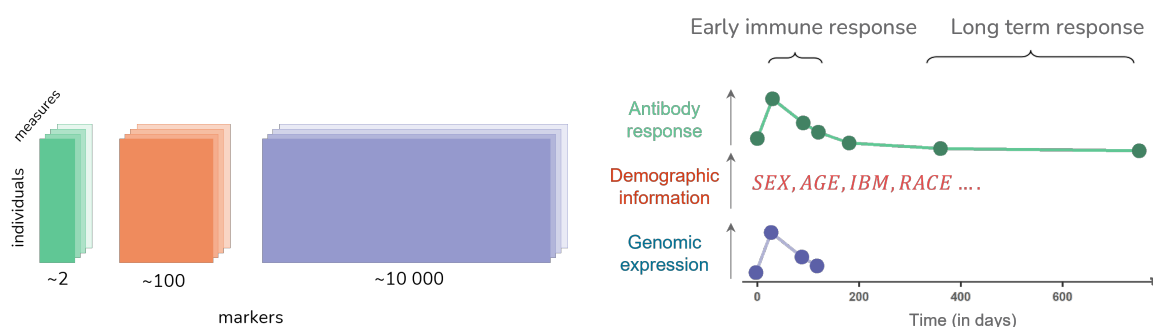


FIGURE 10.1 – Représentation des dimensions courantes pour les données humorales, démographiques et de transcriptomiques.

En adoptant cette approche intégrée, ce projet de recherche contribuera donc à combler une lacune importante dans l'analyse des données transcriptomiques de grande dimension et mécanistiques. En fournissant des outils statistiques avancés pour l'analyse de ces données, cette méthode permettra d'obtenir des résultats plus fiables et interprétables, tout en explorant les relations complexes entre les marqueurs et les processus cellulaires au sein de la réponse immunitaire à la vaccination. Ces résultats auront un impact significatif sur la compréhension de la variabilité individuelle de la réponse aux vaccins



et pourraient potentiellement accélérer le développement de vaccins plus efficaces.

### 10.1.2. Modèles non linéaire à effets mixtes

La dynamique des anticorps joue un rôle essentiel dans la réponse immunitaire. Comprendre les mécanismes biologiques qui contrôlent cette dynamique est donc cruciale pour la recherche médicale et le développement de stratégies vaccinales efficaces. Pour étudier ces phénomènes complexes, nous pouvons utiliser des modèles qui permettent de décrire et d'expliquer les dynamiques cellulaires impliquées [7]. Une approche couramment utilisée pour modéliser ces phénomènes est les modèles à effets mixtes. Les modèles à effets mixtes sont des outils statistiques puissants qui permettent d'analyser les données provenant d'études expérimentales complexes, prenant notamment en compte l'hétérogénéité intra-individuelle présente au sein d'une population. Ils prennent en compte à la fois les facteurs fixes, qui influencent de manière constante les réponses immunitaires humores au sein de population, ou sous-population caractérisée par certaines covariables, mais aussi les effets aléatoires, qui capturent les variations individuelles au sein même d'une sous-population, pour expliquer l'hétérogénéité résiduelle. Ces modèles offrent une approche statistique robuste pour étudier les différences interindividuelles et intra-individuelles dans les réponses immunitaires humores. Ce type de modélisation s'appuie sur des connaissances approfondies des processus biologiques et des interactions moléculaires pour formuler des équations mathématiques décrivant le comportement des systèmes biologiques. Il existe trois types de modèles utilisés au sein d'un modèle mécaniste : le modèle structurelle, le modèle statistique et le modèle d'observation.

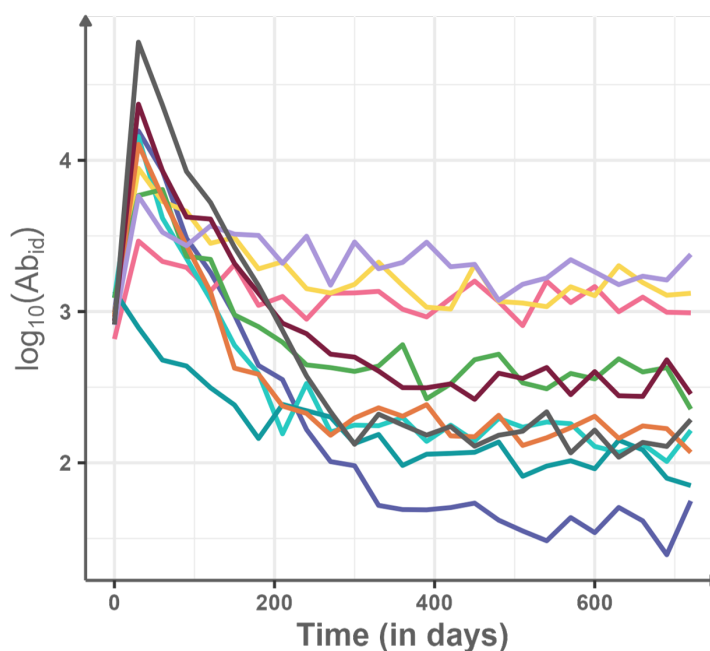


FIGURE 10.2 – Simulations de la réponse immunitaire humorale au vaccin Ebola [9, 1]. On observe des trajectoires bruitées individuelles, et surtout, non linéaire ; les modèles mécanistes offrent la possibilité d'étudier efficacement ces données.

#### Modèle Structurel

Le modèle structurel est le pilier de la modélisation non-linéaire à effets mixtes en biologie. Il représente la dynamique et les interactions entre les composants d'un système biologique. Ces modèles

sont généralement formulés sous forme d'équations différentielles ordinaires (EDO) ou d'équations aux dérivées partielles (EDP), qui décrivent comment les variables biologiques évoluent dans le temps en fonction de l'état actuel du système et des processus régulant leur comportement. Les modèles structuraux peuvent représenter des processus cellulaires, des réseaux de signalisation, des interactions entre protéines, des cascades enzymatiques et bien d'autres phénomènes biologiques complexes. Ils sont conçus pour capturer les mécanismes sous-jacents du système biologique étudié et fournissent une description détaillée de son fonctionnement.

Finalement, si on note  $A_i : \mathbb{R}^+ \rightarrow \mathbb{R}$  la dynamique d'un phénomène biologique pour un individu  $i \leq N$ , que l'on cherche à modéliser. On note  $t_{ij} \in \mathbb{R}^+$  les temps de mesures d'indice  $j \leq n_i$  de  $A_i$ , et on note  $\psi_{ij} \in \mathbb{R}^m$  les paramètres caractérisant le phénomène biologique pour l'individu  $i$  au temps  $t_{ij}$ . On peut alors écrire des modèles de la forme

$$A_i(t = t_{ij}) = f(t_{ij}; \psi_{ij}) \quad \text{ou encore} \quad \frac{d}{dt} A_i(t_{ij}) = f(t_{ij}, A_i; \psi_{ij})$$

où l'indice  $j \leq J$  correspond au indice des temps de mesures pour l'individu.

### Modèle Statistique

Le modèle statistique permet lui de rendre compte de la variabilité intra-individuelle. Il consiste à exprimer les paramètres individuels du modèle en fonction du paramètre moyen et des covariables de l'individus  $i$ . Ces covariables permettent de former des sous-populations qui pourrait présenter des caractéristiques communes. On ajoute à cela des effets aléatoires pour capturer l'hétérogénéité qui ne l'a pas encore été par les covariables. Si donc on note  $X_{ij} \in \mathbb{R}^n$  les covariables au temps  $t_{ij}$  de l'individu  $i$ , on peut écrire, à une transformation  $h : \mathbb{R} \rightarrow \mathbb{R}$  près

$$h(\psi_{ij}) = h(\psi_{pop}) + \beta X_{ij} + \eta_{ij}$$

où  $\eta_{ij} \sim \mathcal{N}(0, \Omega)$  correspond aux effets aléatoires, avec  $\Omega$  la matrice de corrélations des paramètres du modèles.

### Modèle d'Observation

Le modèle d'observation est utilisé pour relier les variables mesurées réellement à celles représentées dans les modèles structuraux. Ces variables observées peuvent être des données expérimentales issues de diverses techniques de mesure, telles que les analyses transcriptomiques, les dosages protéiques, ou les observations cellulaires. Cette partie du modèle permet de traduire les variables biologiques mesurées en termes des processus et des entités biologiques décrits par les modèles structuraux. Il fournit ainsi un lien essentiel entre les données expérimentales et les mécanismes biologiques sous-jacents, permettant une validation et une interprétation plus précises des résultats expérimentaux. Il permet d'inclure notamment les différentes transformations dues aux techniques de mesures, mais aussi le modèle d'erreur. Si on s'intéresse à la dynamique de la fonction  $A_i$ , précédemment introduite, pour un individu  $i \leq N$  au cours du temps. Nos observations  $Y_{ij}$  obtenus de  $A_i(t_{ij})$  seront en réalité des observations bruitées, éventuellement transformée par une fonction  $g : \mathbb{R} \rightarrow \mathbb{R}$ , de sorte que l'on observe en réalité

$$Y_{ij} = g(A_{ij}) + \varepsilon_{ij}$$

avec  $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ , le modèle d'erreur,  $\sigma > 0$ .

### 10.1.3. Pertinence des modèles non-linéaires à effets mixtes en biologie

Les modèles non-linéaires à effets mixtes jouent un rôle crucial dans la modélisation des phénomènes biologiques individuels, car ils permettent de traduire les connaissances biologiques fondamentales en descriptions mathématiques précises et interprétables. Cette approche offre plusieurs avantages clés. D'abord, les modèles non-linéaires à effets mixtes fournissent une représentation détaillée des interactions entre les composants biologiques, ce qui permet de comprendre comment les processus cellulaires et moléculaires se combinent pour donner lieu à des réponses biologiques spécifiques. Les modèles non-linéaires à effets mixtes permettent de relier les données expérimentales aux mécanismes biologiques sous-jacents, offrant une interprétation plus précise des résultats expérimentaux et facilitant l'identification de corrélats biologiques pertinents. Enfin, les modèles non-linéaires à effets mixtes statistiques intègrent des termes d'erreur pour prendre en compte les incertitudes liées aux mesures expérimentales et aux processus biologiques non représentés. Cela permet d'obtenir des estimations plus fiables des paramètres et des prédictions plus précises. En résumé, la modélisation non-linéaire à effets mixtes en biologie est une approche puissante qui permet de relier les connaissances fondamentales des processus biologiques aux données expérimentales observables. Elle offre des possibilités prometteuses pour élucider les mécanismes complexes régissant les réponses biologiques individuelles, ouvrant ainsi de nouvelles perspectives pour la recherche en immunologie et d'autres domaines de la biologie.

### 10.1.4. Modèle à Variables Latentes

Les modèles à variables latentes sont une classe de modèles statistiques qui permettent de représenter des concepts ou des processus sous-jacents non directement observables (les variables latentes) à partir de variables observables. Les variables latentes jouent le rôle de pont entre les données observées et les mécanismes non observables, et leur utilisation est pertinente dans de nombreux domaines scientifiques, y compris la biologie, l'économie, la psychologie, et bien d'autres [11].

Une variable latente est une variable cachée qui n'est pas directement mesurable, mais qui influence les variables observables que l'on peut mesurer. Les variables latentes sont utilisées pour représenter des concepts ou des phénomènes qui ne sont pas directement accessibles par les mesures expérimentales ou les observations, mais qui sont supposés jouer un rôle crucial dans la génération des données observées. Les modèles à variables latentes présentent plusieurs avantages et intérêts dans l'analyse des données, notamment en biologie :

1. Les variables latentes permettent de modéliser des mécanismes biologiques sous-jacents qui ne sont pas directement observables. Par exemple, dans le contexte de la réponse immunitaire à la vaccination, certaines réponses cellulaires ou métaboliques peuvent être des variables latentes qui influencent les marqueurs transcriptomiques observables.
2. Les modèles à variables latentes peuvent être utilisés pour réduire la dimension des données en résumant l'information contenue dans plusieurs variables observables en une ou plusieurs variables latentes. Cela permet de simplifier la complexité des données tout en conservant l'essentiel des informations.
3. Les variables latentes peuvent aider à gérer l'incertitude et les erreurs de mesure dans les données. En prenant en compte les relations entre les variables observables et les variables latentes, les modèles à variables latentes permettent de modéliser les erreurs de mesure et d'estimer les valeurs réelles des variables latentes malgré le bruit dans les données.

4. Les variables latentes fournissent une représentation plus abstraite et conceptuelle des données, ce qui facilite l'interprétation des résultats et la compréhension des phénomènes sous-jacents. Cela peut conduire à des découvertes importantes sur les mécanismes biologiques qui ne seraient pas facilement apparents sans l'utilisation de variables latentes.

Dans les modèles non-linéaires à effets mixtes basés sur des équations différentielles ordinaires (EDO) ou d'autres approches similaires, la modélisation de nombreux marqueurs transcriptomiques peut rapidement devenir complexe et conduire à des modèles surchargés. Les variables latentes agissent comme des variables agrégées qui résument l'information contenue dans plusieurs marqueurs observables. En regroupant les marqueurs en variables latentes, on réduit considérablement la dimensionnalité du modèle, ce qui facilite la spécification et l'interprétation du modèle non-linéaire à effets mixtes. Les modèles à variables latentes sont un outil puissant pour la modélisation des phénomènes biologiques complexes, en permettant de représenter des mécanismes cachés, de réduire la dimensionnalité des données, de gérer l'incertitude et d'améliorer l'interprétation des résultats. Ces modèles constituent une approche polyvalente et pertinente dans l'analyse des données biologiques, en particulier dans le cadre du projet de recherche sur la réponse immunitaire à la vaccination, où les processus cellulaires et métaboliques sous-jacents peuvent être représentés par des variables latentes pour une meilleure compréhension des mécanismes impliqués. Leur dans des modèles mécanistes pour la grande dimension offre une approche puissante et pertinente pour gérer la complexité et l'hétérogénéité des données biologiques tout en capturant les mécanismes cachés. Cette approche permet également de réduire le risque de sur-ajustement et d'intégrer les incertitudes, tout en facilitant l'interprétation des résultats pour une meilleure compréhension de la réponse immunitaire à la vaccination.

#### 10.1.5. L'exemple de la réponse immunitaire humorale

Prenons l'exemple de la réponse immunitaire humorale représenté en figure 3. Les anticorps  $Ab$  sont sécrétés respectivement aux taux  $\theta_S$  et  $\theta_L$  par des cellules sécrétrices d'anticorps  $S$  et  $L$ , caractérisés par leur durée de vie. Les cellules  $S$  pour « short-live » ont une demi-vie moyenne de 3 jours et celles  $L$  pour « long-live » ont une demi-vie moyenne estimée à 6 ans. Chacune de ces entités biologiques se désintègre respectivement au taux  $\delta_{Ab}, \delta_S$  et  $\delta_L$ .

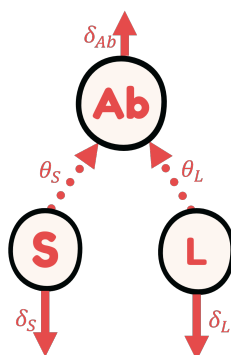


FIGURE 10.3 – Schéma de la production d'anticorps dans le cadre d'une réponse immunitaire humorale [9]

La dynamique peut donc se résumer à cette équation différentielle :

$$\frac{d}{dt}Ab(t) = \underbrace{S_0\theta_S}_{\varphi_S} e^{-\delta_S t} + \underbrace{L_0\theta_L}_{\varphi_L} e^{-\delta_L t} - \delta_{Ab} Ab(t)$$

Le modèle non-linéaire à effets mixtes utilisé pour modéliser cette dynamique est :

►► Modèle Structurel :

$$\begin{aligned} \forall i \leq N, j \leq n_i, \quad & \begin{cases} \frac{d}{dt}Ab_i(t_{ij}) &= \varphi_{S_i} e^{-\delta_S t_{ij}} + \varphi_{L_i} e^{-\delta_L t_{ij}} - \delta_{Ab_i} Ab_i(t_{ij}) \\ Ab_i(t_{i0} = 0) &= Ab_{0i} \end{cases} \\ & \begin{cases} \log(\varphi_{S_i}) &= \log(\varphi_{S_{pop}}) + \log(FC\Delta AGE = +1an) \times (AGE_i - \overline{AGE}) + \eta_i^S \\ \log(\varphi_{L_i}) &= \log(\varphi_{L_{pop}}) + \beta_{RACE} \times 1_{RACE_i=European} + \eta_i^L \\ \log(\delta_{Ab_i}) &= \log(\delta_{Ab_{pop}}) + \beta_{SEX} \times 1_{SEX_i=Male} + \eta_i^{Ab} \end{cases} \end{aligned} \quad (\text{MOD AB STR})$$

►► Modèle Statistique :

$$\forall i \leq N, \quad \begin{cases} \eta_i^S &\stackrel{iid}{\sim} \mathcal{N}(0, \omega_S^2) \\ \eta_i^L &\stackrel{iid}{\sim} \mathcal{N}(0, \omega_L^2) \\ \eta_i^{Ab} &\stackrel{iid}{\sim} \mathcal{N}(0, \omega_{Ab}^2) \end{cases} \quad (\text{MOD AB STAT})$$

►► Modèle d'Observation :

$$\forall i \leq N, j \leq n_i, \quad \begin{cases} Y_{ij} &= \log_{10}(Ab_i(t_{ij})) + \varepsilon_{ij} \\ \varepsilon_i &\stackrel{iid}{\sim} \mathcal{N}(0, \Sigma = \sigma_{Ab}^2 I_{n_i}) \end{cases} \quad (\text{MOD AB OBS})$$

(MOD AB)

Dans cet exemple, la variable observé ici est la quantité d'anticorps chez l'individu  $i \leq N$ . La dynamique des populations de cellules sécrétrices d'anticorps elle n'est pas observée. Cependant, celle-ci pourrait être inférée en analysant les données de génomiques recueillies. Comprendre mieux l'influence et le comportement de ces cellules sur le système immunitaire apporterait de nouvelles perspectives pertinentes pour mieux comprendre la réponse immunitaire.

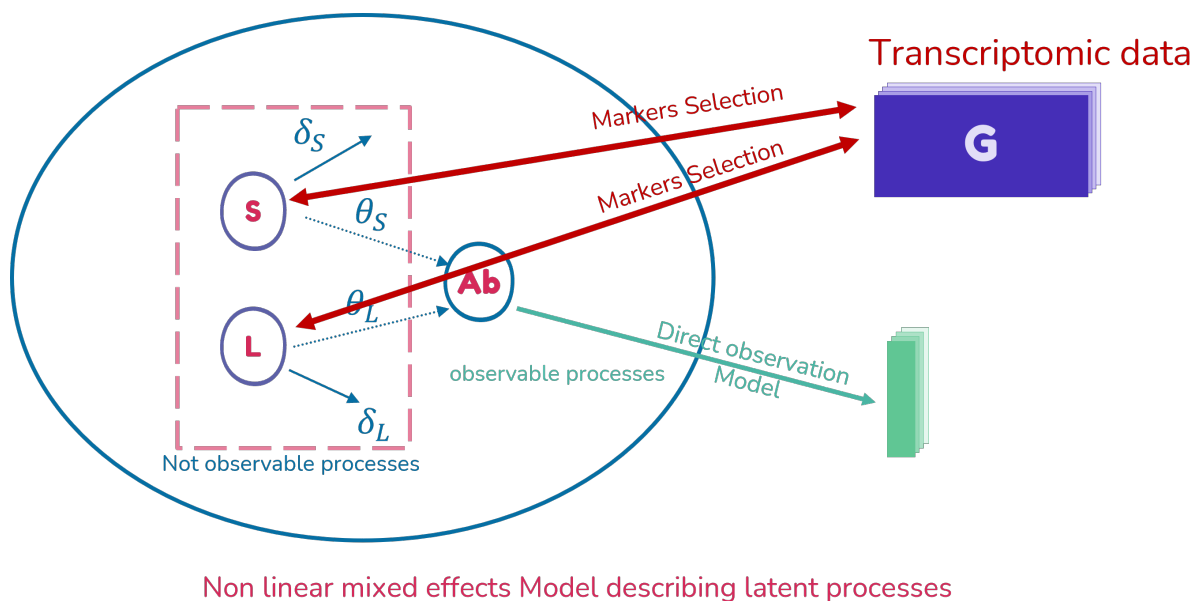


FIGURE 10.4 – Schéma du processus latent pour le modèle non-linéaire à effets mixtes de la réponse immunitaire.

## 10.2. Analyse Méthodologique et Travail envisagé

Dans le but d'incorporer des données de grandes dimensions dans les modèles non-linéaires à effets mixtes, deux approches sont possible. Une première qui consiste à modifier le modèle statistique en effectuant une sélection de variables. Ceci est l'objet du stage réalisé en amont de la thèse dans la même équipe. Une seconde approche est l'utilisation des modèles à variables latents au sein donc du modèle d'observation.

### 10.2.1. Travail sur le modèle statistique

En amont de la thèse, je réalise un stage au sein de l'équipe SISTM, encadré par Mélanie Prague sur la problématique de sélection de covariables pour les modèles non-linéaires à effets mixtes. Ce stage a permis de me familiariser avec tous les concepts nécessaires à la bonne réalisation de ce projet de doctorat, mais aussi de m'intégrer efficacement dans l'équipe et de prendre mes marques avant le début de thèse. Cette approche consiste donc à sélectionner les covariables dans le modèle statistique décrit précédemment. Il existe à l'heure actuellement un algorithme de construction de modèle, SAMBA [10] par Mélanie Prague et Marc Lavielle. Cette algorithme est performant pour construire à la fois le modèle

statistique et d'erreur d'un modèle non-linéaire à effets mixtes . Il détermine le modèle statistique en sélectionnant les covariables significatives par méthode stepwise sur le critère d'information et en estimant la matrice de corrélations  $\Omega$ . Ces méthodes ne sont pas robustes à la grande dimension et l'objet du stage est donc d'incorporer les méthodes de regression par lasso et d'apprentissage statistique pour combler ce manque. En faisant de la sorte, on ne réduit cependant pas le risque de propagation de l'incertitude.

### 10.2.2. Travail sur le modèle d'observation

Pour le projet de thèse, deux approches sont envisagées pour incorporer les modèles à variables latentes dans les modèles non-linéaires à effets mixtes. Si on reprends notre exemple précédent, en figure 4., les variables latentes sont les populations S et L de cellules sécrétrices d'anticorps. Nos données de génétiques, noté  $G_{ij}^k$  pour le gènes  $k \leq K$ , de l'individu  $i \leq N$  au temps  $t_{ij}$ , peuvent alors être utilisé pour inférer la dynamique de ces sous-populations cellulaires.

#### Approche Réflexive

Une première approche pour le faire serait d'utiliser une approche réflexive pour le modèle d'observation. Cela signifie que l'on écrirait le modèle d'observation sous la forme :

$$\forall k \leq K, i \leq N, \quad G_{ij}^k = h_k(S_{ij}) + \varepsilon_{ij}^k$$

Les expressions génétiques mesurées sont supposées être une transformation bruitée de la population de cellules sécrétrices d'anticorps – la population S ici dans l'exemple.

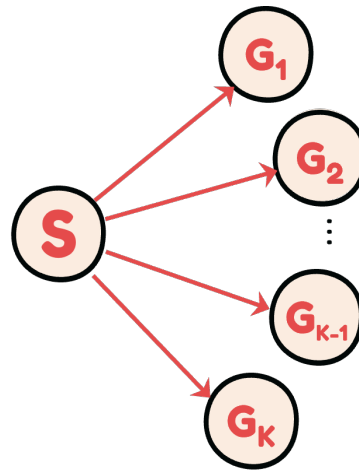


FIGURE 10.5 – Schéma réflexif du modèle d'observation

#### Approche Formative

Une seconde approche pour le faire serait d'utiliser une approche formative pour le modèle d'observation. Cela signifie que l'on écrirait le modèle d'observation sous la forme :

$$\forall i \leq N, \quad S_{ij} = h_k(G_{ij}^1 + \varepsilon_{ij}^1, \dots, G_{ij}^K + \varepsilon_{ij}^K)$$

Les expressions génétiques mesurées, et donc bruitées, sont supposées déterminer la dynamique de la population de cellules sécrétrices d'anticorps.

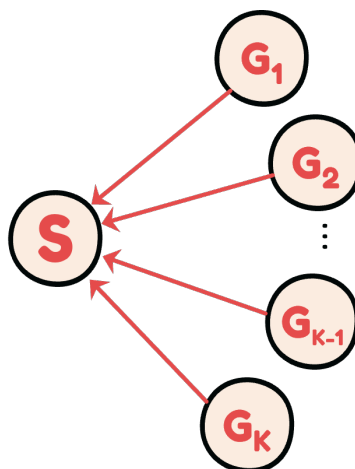


FIGURE 10.6 – Schéma formatif du modèle d'observations

### Pertinence et Contexte

Le projet pour ces deux approches est de les étudier toute deux et d'en avoir recours selon le contexte structurel du phénomène biologique. L'approche la plus adapté aux hypothèses biologiques et cliniques serait utilisé. Dans un premier temps, le projet de thèse se concentrera sur des transformations reliant les variables latentes à celle observées indépendantes du temps. Elles seront ensuite possiblement dépendante du temps, et les modèles s'écrirait sous la forme :

$$\begin{array}{l|l} \text{Approche Réflective} & \forall k \leq K, i \leq N, \quad G_{ij}^k = h_k(t_{ij}; S_{ij}) + \varepsilon_{ij}^k \\ \text{Approche Formative} & \forall i \leq N, \quad S_{ij} = h_k(t_{ij}; G_{ij}^1 + \varepsilon_{ij}^1, \dots, G_{ij}^K + \varepsilon_{ij}^K) \end{array}$$

### 10.2.3. L'étude Prevac-UP, et autres applications

Tout au long de cette thèse, de nombreuses applications pratiques et concrètes seront possibles grâce à la méthode développée. Les données disponibles provenant de divers projets de vaccination, tels que PREVAC-UP [12] et Ebovac [8] (European IMI2 projects) pour Ebola, ainsi que les collaborations avec le Vaccine Research Institute (VRI), offriront un terrain fertile pour l'application et la validation des résultats obtenus. En utilisant cette nouvelle approche intégrée de modélisation non-linéaire à effets mixtes à variables latentes, il sera possible d'analyser ces données de grande dimension avec une plus grande précision et une meilleure interprétation, ouvrant ainsi la voie à de nouvelles découvertes et avancées dans le domaine de la vaccination. Par exemple, le projet PREVAC-UP sera un cas d'application intéressant pour étudier les réponses immunitaires individuelles et identifier les corrélats de protection potentiels dans le cas du vaccin contre Ebola. De même, les collaborations avec le Vaccine Research Institute pour la vaccination contre le SARS-CoV-2 permettront d'approfondir notre compréhension de la variabilité individuelle de la réponse aux vaccins et de mieux cibler les vaccins prometteurs pour accélérer leur développement. Les résultats de cette thèse auront également un impact important sur d'autres projets européens dans lesquels l'équipe est impliquée en tant que Work Package Leaders, tels que l'European HIV Vaccine Alliance (EHVA) pour le VIH. Ces projets pourraient également bénéficier de l'utilisation de



la méthode développée pour améliorer l'analyse des données de grande dimension et mieux comprendre les mécanismes sous-jacents aux réponses immunitaires. Enfin, cette approche intégrée de modélisation non-linéaire à effets mixtes à variables latentes pourrait également avoir un intérêt pour des collaborateurs extérieurs, tels que les équipes [MONC](#) ou [CARMEN](#), qui font face à des problèmes similaires de modélisation de données de grande dimension dans d'autres domaines scientifiques. Ainsi, cette thèse pourrait ouvrir de nouvelles perspectives de collaboration et d'application dans divers domaines de la recherche en biologie et en sciences de la santé.



FIGURE 10.7 – A gauche logo du projet Prevac-UP, essai vaccinale contre Ebola ; A droite, logo du Vaccine Research Institute (VRI), avec qui l'équipe SISTM travaille.

---

## 10.3. AlkenSus

---

### 10.3.1. Calendrier des trois années de thèse

Tout d'abord, un stage réalisée du 1er avril 2023 au 30 septembre 2023 préparera la thèse. Le sujet connexe permettra de se familiariser avec les notions nécessaires et les outils et packages utilisés. Le calendrier prévisionnel de cette thèse se déroulera sur une période de trois ans. Le premier semestre pourra se consacrer à la rédaction d'un article sur les méthodes implémenté pendant le stage et en vue de la thèse. En même temps, le premier projet de la thèse sera lancer, sur l'incorporation des deux approches cités dans des modèles mécanistes avec solution analytique et sans dépendance de temps entre les variables observées et celles latentes. Cette phase initiale permettra de poser les bases théoriques nécessaires à la mise en œuvre du projet de recherche et de le débiter avec des hypothèses plus souple. La seconde partie relâchera les hypothèses faites sur le modèle non-linéaire à effets mixtes pour tendre vers un cadre d'application plus large. Les deux approches cités seront encore étudié, sans préférence à priori. Au cours de la thèse, et selon la fiabilité des hypothèses, les approches pourront être modifié. Tout au long, des simulations et des applications sur données réelles seront réalisées pour évaluer la performance de la méthode et comparer les résultats obtenus avec d'autres approches existantes. Les applications sur données réelles seront disponibles tout au long de la thèse grâce au travail existant entre l'équipe SISTM et le VRI par exemple.

### Use of latent variable models to build mechanistic models in high dimension

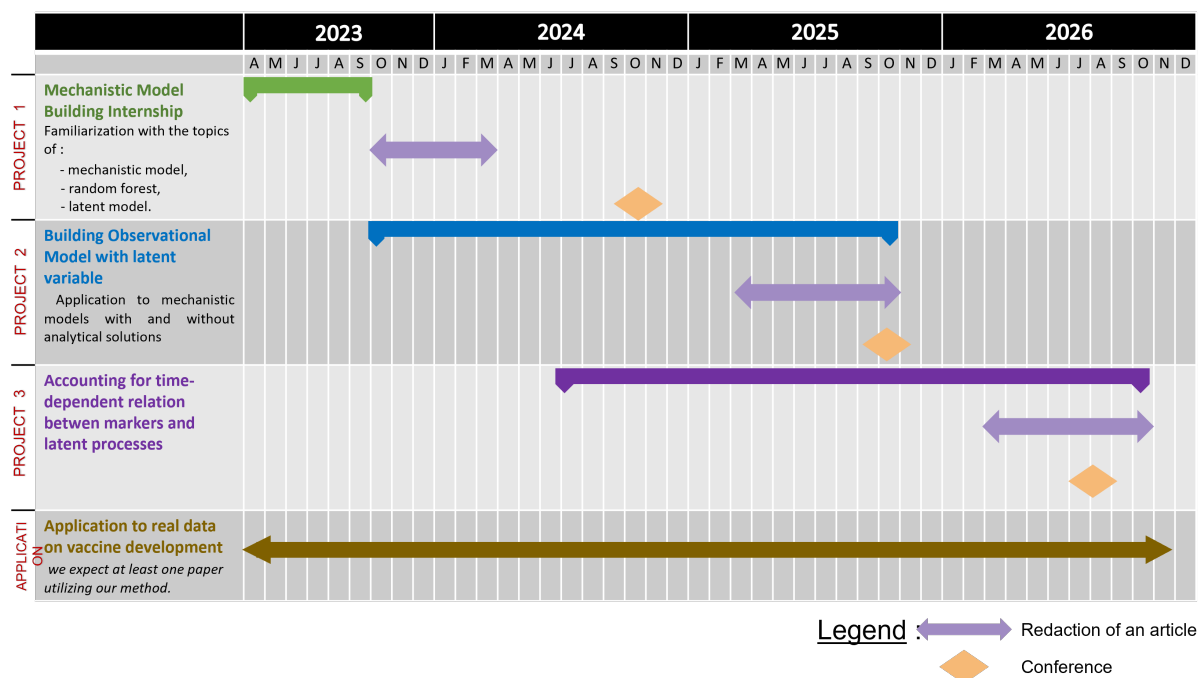


FIGURE 10.8 – Gant du calendrier prévisionnel pour le projet de thèse.

Il est attendu au cours de cette thèse la rédaction de 3 articles dont 1 publiés, ainsi que des participations à des conférences internationales. Enfin, le dernier semestre sera dédié à la rédaction de la thèse, à la préparation de la soutenance et à la dissémination des résultats à travers des publications, des présentations et la mise à disposition du code source via Github et une éventuelle création d'un package R. Le calendrier prévisionnel de cette thèse vise à atteindre les objectifs fixés tout en étant flexible pour s'adapter aux éventuels imprévus et ajustements nécessaires pour mener à bien ce projet de recherche. Le calendrier est représenté sur le schéma figure 6.

A la date d'écriture de ce mémoire, une présentation de poster sur le sujet de stage, lié à la thèse, a été réalisé le 18 septembre 2023 à l'occasion de la journée des jeunes chercheurs du BPH (Bordeaux Population Health research center). Cette présentation a été récompensé du prix du "meilleur poster" et de la somme de 500€ pour une conférence. Un abstract à également était soumis pour la journée de Biostatistique à l'occasion des GDR de Toulouse. Un aperçu du poster évoqué et l'abstract soumis son ajouté en annexe de ce mémoire.

#### 10.3.2. Perspectives

Le présent projet de recherche vise à développer une méthodologie innovante et intégrée pour l'analyse de données de grande dimension dans le contexte de la modélisation non-linéaire à effets mixtes de la réponse immunitaire à la vaccination. Les résultats attendus incluent la mise au point d'une méthode robuste pour la sélection de marqueurs transcriptomiques pertinents et leur intégration dans des modèles mécanistes. Cette approche permettra de mieux comprendre les mécanismes sous-jacents qui régulent la réponse immunitaire individuelle, en identifiant directement les associations entre les marqueurs et les processus cellulaires impliqués. Ce projet de recherche, en fournissant une approche intégrée pour l'ana-

lyse des données de grande dimension, permettra d'éviter la propagation d'incertitudes causée par une sélection de marqueurs en amont, améliorant ainsi la précision et la fiabilité des résultats. Cela sera d'une importance capitale pour la recherche en immunologie et la compréhension de la variabilité individuelle dans la réponse aux vaccins. De plus, cette approche novatrice permettra de mieux exploiter les données de transcriptomique de grande dimension, en fournissant des informations riches sur les mécanismes cellulaires sous-jacents. En identifiant les marqueurs les plus pertinents, cette méthode pourra contribuer à la découverte de corrélats biologiques précoces de protection vaccinale, ce qui pourrait accélérer le développement de vaccins plus efficaces. Les perspectives découlant de ce projet de recherche sont multiples et prometteuses. Tout d'abord, cette méthodologie pourra être appliquée à d'autres domaines de la biologie et de la médecine où des données de grande dimension sont disponibles, élargissant ainsi son champ d'application. En outre, en améliorant la compréhension de la variabilité individuelle dans la réponse immunitaire, cette approche pourrait conduire à des avancées importantes dans la médecine personnalisée, en permettant une meilleure adaptation des stratégies vaccinales en fonction des caractéristiques individuelles. Enfin, cette méthodologie pourrait également susciter l'intérêt de la communauté scientifique internationale et stimuler la collaboration avec d'autres équipes de recherche travaillant sur des problèmes similaires. La diffusion des résultats à travers des publications et des conférences pourrait favoriser les échanges scientifiques et ouvrir de nouvelles perspectives de recherche dans le domaine de la modélisation non-linéaire à effets mixtes pour la grande dimension. En somme, ce projet de recherche offre une opportunité unique d'allier la modélisation non-linéaire à effets mixtes avancée avec les défis actuels de l'analyse des données de grande dimension, contribuant ainsi à la compréhension plus approfondie de la réponse immunitaire à la vaccination et ouvrant de nouvelles perspectives pour le développement de vaccins plus efficaces et personnalisés.

---

## Liste des références

---

ALEXANDRE, M., PRAGUE, M., MCLEAN, C., BOCKSTAL, V., DOUGUIH, M. et THIÉBAUT, R. (2023). Evaluation and prediction of the long-term humoral immune response induced by the two-dose heterologous ad26.zebov,mva-bn-filo vaccine regimen against ebola.

BALELLI, I., PASIN, C., PRAGUE, M., CRAUSTE, F., EFFELTERRE, T. V., BOCKSTAL, V., SOLFOROSI, L. et THIÉBAUT, R. (2020). A model for establishment, maintenance and reactivation of the immune response after vaccination against ebola virus. *Journal of Theoretical Biology*, 495 :110254.

CAPITAINE, L., GENUER, R. et THIÉBAUT, R. (2021). Random forests for high-dimensional longitudinal data. *Statistical Methods in Medical Research*, 30(1) :166–184. PMID : 32772626.

CLAIRON, Q., PRAGUE, M., PLANAS, D., BRUEL, T., HOCQUELOUX, L., PRAZUCK, T., SCHWARTZ, O., THIÉBAUT, R. et GUEDJ, J. (2023). Modeling the kinetics of the neutralizing antibody response against sars-cov-2 variants after several administrations of bnt162b2. *PLOS Computational Biology*, 19 :e1011282.

DUVIGNAUD, A., LHOMME, E., PISTONE, T., ONAISI, R., SITTA, R., JOURNOT, V., NGUYEN, D., PEIFFER-SMADJA, N., CRÉMER, A., BOUCHET, S., DARNAUD, T., POITRENAUD, D., PIROTH, L., BIN-

QUET, C., MICHEL, J. F., LEFÈVRE, B., LEBEAUX, D., LEBEL, J., DUPOUY, J., ROUSSILLON, C., GIMBERT, A., WITTKOP, L., THIÉBAUT, R., ORNE-GLIEMANN, J., JOSEPH, J. P., RICHERT, L., ANGLARET, X. et MALVY, D. (2020). Home treatment of older people with symptomatic sars-cov-2 infection (covid-19) : A structured summary of a study protocol for a multi-arm multi-stage (mams) randomized trial to evaluate the efficacy and tolerability of several experimental treatments to reduce the risk of hospitalisation or death in outpatients aged 65 years or older (coverage trial).

GAUTHIER, M., AGNIEL, D., THIÉBAUT, R. et HEJBLUM, B. P. (2020). Dearseq : A variance component score test for rna-seq differential analysis that effectively controls the false discovery rate. *NAR Genomics and Bioinformatics*, 2.

LAVIELLE, M. (2014). *Mixed Effects Models for the Population Approach*. Chapman and Hall/CRC.

MOONEY, T., SMOUT, E., LEIGH, B., GREENWOOD, B., ENRIA, L., ISHOLA, D., MANNO, D., SAMAI, M., DOUGUIH, M. et WATSON-JONES, D. (2018). Ebovac-salone : Lessons learned from implementing an ebola vaccine trial in an ebola-affected country. *Clinical Trials*, 15 :436–443.

PASIN, C., BALELLI, I., EFFELTERRE, T. V., BOCKSTAL, V., SOLFOROSI, L., PRAGUE, M., DOUGUIH, M. et THIÉBAUT, R. (2019). Dynamics of the humoral immune response to a prime-boost ebola vaccine : Quantification and sources of variation. *Journal of Virology*, 93.

PRAGUE, M. et LAVIELLE, M. (2022). Samba : A novel method for fast automatic model building in nonlinear mixed-effects models. *CPT : Pharmacometrics & Systems Pharmacology*, 11 :161–172.

PROUST-LIMA, C., AMIEVA, H. et JACQMIN-GADDA, H. (2013). Analysis of multivariate mixed longitudinal data : A flexible latent process approach. *British Journal of Mathematical and Statistical Psychology*, 66 :470–487.

TEAM, P. S. (2022). Randomized trial of vaccines for zaire ebola virus disease. *New England Journal of Medicine*, 387 :2411–2424.

Annexe

,max size=0.9,widthh=

## A. Cours en classe de MPSI

**NOM/Prénom :** \_\_\_\_\_ le 26/05/2021

**Interrogation écrite n°22**

**Cours :**

Base canonique de  $\mathcal{M}_{p,n}(\mathbb{K})$  : \_\_\_\_\_

$\dim(D_n(\mathbb{K})) = \underline{\hspace{2cm}}$  et  $\dim(D_n^1(\mathbb{K})) = \underline{\hspace{2cm}}$

Si  $A = (a_{ij}) \in \mathcal{M}_{p,n}(\mathbb{K})$  et  $B = (b_{ij}) \in \mathcal{M}_{r,n}(\mathbb{K})$ ,  $A \times B$  existe ssi \_\_\_\_\_ et  $A \times B = (p_{ij}) \in \mathcal{M}_{\underline{\hspace{1cm}}}(\mathbb{K})$  avec :  $\forall (i,j) \in \llbracket 1, \underline{\hspace{1cm}} \rrbracket \times \llbracket 1, \underline{\hspace{1cm}} \rrbracket : p_{ij} = \underline{\hspace{2cm}}$

Si  $A \in \mathcal{M}_n(\mathbb{K})$ , on dit que  $A$  est nilpotente ssi \_\_\_\_\_

$A, B \in \mathcal{M}_n(\mathbb{K})$  tels que \_\_\_\_\_ alors  $p \in \mathbb{N}^*$  :  $A^p \cdot B^p = \underline{\hspace{2cm}}$

**Exercice :**

On choisit 2 puis 4 lettres au hasard une à une dans le mot ANANAS, quelle est la probabilité d'obtenir dans cet ordre, le mot AN ? le mot NANA ?

FIGURE 9 – Interrogation écrite n°22

**NOM/Prénom :** \_\_\_\_\_ le 26/05/2021

**Interrogation écrite n°23**

**Cours :**

Matrice de passage de  $B$  vers  $B'$  :  $P^B_{B'} = \underline{\hspace{2cm}}$

**Schéma des matrices de passage :** \_\_\_\_\_

**Formules de changement de bases :** \_\_\_\_\_

**Vecteurs :** \_\_\_\_\_

**Applications linéaires :** \_\_\_\_\_

**Endomorphismes :** \_\_\_\_\_

**Exercice :**

On considère  $E = \mathbb{R}_2[X]$  et  $\varphi : \begin{cases} E \rightarrow E \\ P(X) \mapsto P(X) + (X^2 + 1)P(0) + XP(1) \end{cases}$

- 1) Montrer que  $\varphi$  est un endomorphisme de  $E$ .
- 2) Donner la matrice de  $\varphi$  dans la base canonique de  $E$ .
- 3) Montrer que  $\varphi$  est un automorphisme et donner  $\varphi^{-1}$ .

FIGURE 10 – Interrogation écrite n°23

NOM/Prénom : \_\_\_\_\_ le 26/05/2021

Interrogation écrite n°24

Cours :

Si A et B de  $\mathcal{M}_{p,n}(\mathbb{K})$ , A et B sont dites équivalentes ssi \_\_\_\_\_ (def)

Si A et B de  $\mathcal{M}_n(\mathbb{K})$ , A et B sont dites semblables ssi \_\_\_\_\_ (caract)

Si A et B de  $\mathcal{M}_n(\mathbb{K})$ , A et B sont dites semblables ssi \_\_\_\_\_ (def)

Si A et B de  $\mathcal{M}_n(\mathbb{K})$ , A et B sont dites semblables ssi \_\_\_\_\_ (caract)

Si A = (a<sub>ij</sub>)<sub>1 ≤ i,j ≤ n</sub> ∈  $\mathcal{M}_n(\mathbb{K})$ , on appelle trace de A : Tr(A) = \_\_\_\_\_

Propriétés de la trace :

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Exercice :

Calcul de l'inverse de A =  $\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}$  par la méthode du pivot de Gauss.

FIGURE 11 – Interrogation écrite n°24

NOM/Prénom : \_\_\_\_\_ le 07/06/2021

Interrogation écrite n°25

Cours :

Si X et Y deux variables aléatoires sur (Ω, P), ON dit que X et Y sont indépendantes ssi \_\_\_\_\_

Lemme des coalitions : \_\_\_\_\_

Théorème de transfert : \_\_\_\_\_

Si X ~  $\mathcal{U}([1, n])$ , alors E(X) = \_\_\_\_\_ et V(X) = \_\_\_\_\_

Si X ~  $\mathcal{B}(n, p)$ , alors E(X) = \_\_\_\_\_ et V(X) = \_\_\_\_\_

Cov(X, Y) = \_\_\_\_\_ et V(X+Y) = \_\_\_\_\_

Exercice :

X est une variable aléatoire réelle sur Ω fini à valeurs dans  $\llbracket 0, N \rrbracket$ , montrer que  $E(X) = \sum_{k=0}^{N-1} P(X > k)$ .

FIGURE 12 – Interrogation écrite n°25

Classes de MPSI1 et 2  
Mathématiques

Année 2020-2021.  
à rendre pour le 26/05/2021.

**Devoir à la maison n°21.**

**Problème :**

Une particule se déplace sur un triangle  $ABC$ , à l'instant  $t=0$ , la particule se trouve en  $A$ .

Ensuite lorsque la particule est à une position à l'instant  $n$  ( $n \in \mathbb{N}$ ), elle quitte sa position et se déplace de manière équiprobable à l'instant  $n+1$  sur l'une des deux positions non occupées.

On note pour  $n \in \mathbb{N}$ ,  $a_n$ ,  $b_n$  et  $c_n$  les probabilités pour que la particule soit en  $A$ ,  $B$  ou  $C$  à l'instant  $n$ .

Le problème a pour but de calculer les puissances d'une matrice par différentes méthodes, chaque méthode est indépendante de la précédente, les résultats des questions 4) et 5) ne doivent pas utiliser les résultats de la question 2).

- 1) a) Calculer  $a_0$ ,  $b_0$ ,  $c_0$ .  
b) Exprimer pour  $n \in \mathbb{N}$ ,  $a_{n+1}$ ,  $b_{n+1}$  et  $c_{n+1}$  en fonction de  $a_n$ ,  $b_n$  et  $c_n$ .
- 2) On considère  $A = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$  et  $B = 2A + I_3$ .  
a) Calculer pour tout  $n \in \mathbb{N}$ ,  $B^n$ .  
b) En déduire par la formule du binôme de Newton l'expression de  $A^n$  pour tout  $n \in \mathbb{N}$ .  
c) En déduire l'expression de  $a_n$ ,  $b_n$  et  $c_n$  pour tout  $n \in \mathbb{N}$ .  
d) Que se passera-t-il à long terme pour la particule ?
- 3) a) Montrer que  $A$  est inversible et calculer  $A^{-1}$ .  
b) L'expression de  $A^n$  trouvée en 2)b) est-elle valable pour  $n \in \mathbb{Z}$  ?
- 4) a) Montrer que  $2A^2 = A + I_3$ .  
b) Montrer qu'il existe deux suites  $(\beta_n)$  et  $(\gamma_n) \in \mathbb{R}^{\mathbb{N}}$  telles que pour tout  $n \in \mathbb{N}$  on ait  $A^n = \beta_n A + \gamma_n I_3$  et déterminer une relation de récurrence entre  $\beta_{n+1}$  et  $\beta_n$ ,  $\gamma_{n+1}$  et  $\beta_n$ ,  $\gamma_n$ .  
c) Donner l'expression de  $\beta_{n+2}$  en fonction de  $\beta_{n+1}$  et  $\beta_n$  pour tout  $n \in \mathbb{N}$ .  
d) Déterminer alors l'expression pour tout  $n \in \mathbb{N}$  des suites  $(\beta_n)$  et  $(\gamma_n)$ .  
e) Retrouver ainsi le résultat de la question 2)b).
- 5) On note  $a$  l'endomorphisme de  $\mathbb{R}^3$  canoniquement associé à  $A$  et  $\text{id}$  l'endomorphisme identité de  $\mathbb{R}^3$ .  
a) Montrer que  $a$  est un automorphisme de  $\mathbb{R}^3$ .  
b) Déterminer les réels  $\alpha$  tels que  $a - \alpha \text{id}$  n'est pas bijective.  
c) Pour chaque réel  $\alpha$  trouvé en 5)b) déterminer une base de  $\text{Ker}(a - \alpha \text{id})$ .  
d) En déduire une matrice inversible  $P \in \text{GL}_3(\mathbb{R})$  telle que  $P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$ .  
e) En déduire l'expression de  $A^n$  en fonction de  $P$  et  $D$  pour  $n \in \mathbb{N}$ .  
f) Retrouver ainsi le résultat de la question 2)b).
- 6) On considère le polynôme  $B(X) = (X-1)\left(X + \frac{1}{2}\right)$ , on note  $R_n(X)$  le reste de la division euclidienne de  $X^n$  par  $B(X)$  pour  $n \in \mathbb{N}$ .  
a) Déterminer  $R_n(X)$  pour tout  $n \in \mathbb{N}$ .  
b) En notant  $R_n(X) = \lambda_n X + \mu_n$ , montrer que pour tout  $n \in \mathbb{N}$  :  $A^n = \lambda_n A + \mu_n I_3$ .

FIGURE 13 – Enoncé du DM



Classe de MPSI 2  
Mathématiques

Année 2020-2021.  
Semaine du 25/05 au 29/05.

**Programme de colle n°28**

**Matrices :**

- Matrices à  $p$  lignes et  $n$  colonnes à coefficients dans  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ . Vocabulaire matriciel.
- L'espace vectoriel  $\mathcal{M}_{p,n}(\mathbb{K})$ , base canonique, dimension, sous espaces des matrices diagonales, triangulaires supérieures (resp inférieures).
- Produit matriciel : définition et propriétés.  
L'anneau  $\mathcal{M}_n(\mathbb{K})$ , non commutativité pour  $n \geq 2$ , puissances de matrices, matrices nilpotentes, matrices qui commutent, formule du binôme et factorisation  $A^n - B^n$  pour deux matrices carrées qui commutent. Le sous-anneau des matrices diagonales et des matrices triangulaires supérieures (resp inférieures).
- Matrices carrées inversibles : définition,  $I_n$  est inversible, si  $A$  est inversible alors  $A^{-1}$  est inversible et  $(A^{-1})^{-1} = A$ , si  $A$  et  $B$  sont inversibles alors  $A \times B$  est inversible et expression de  $(A \times B)^{-1}$ .  
Une matrice inversible est régulière à gauche et à droite.  
Une matrice carrée est inversible ssi elle est inversible à gauche ssi elle est inversible à droite.
- Rang d'une matrice : c'est le rang de ses vecteurs colonnes. Si  $A \in \mathcal{M}_{p,n}(\mathbb{K})$  alors  $\text{rg}(A) \leq \min(n, p)$ ,  $\text{rg}(A)$  ssi les  $n$  colonnes de  $A$  forment une famille libre de  $\mathbb{K}^p$  et  $\text{rg}(A) = p$  ssi les  $n$  colonnes de  $A$  forment une famille génératrice de  $\mathbb{K}^p$ .
- Transposée d'une matrice : Définition, propriétés : linéarité, transposée d'un produit, transposée d'une matrice inversible. La transposition est un isomorphisme de  $\mathcal{M}_{p,n}(\mathbb{K})$  dans  $\mathcal{M}_{n,p}(\mathbb{K})$  et un automorphisme de  $\mathcal{M}_n(\mathbb{K})$ . Matrices symétriques et antisymétriques : définitions, notation,  $\mathcal{S}_n(\mathbb{K})$  et  $\mathcal{A}_n(\mathbb{K})$  sont des sev supplémentaires de  $\mathcal{M}_n(\mathbb{K})$ , dimension de ces sous-espaces.
- Si  $E$  est un  $\mathbb{K}$ -ev de dimension finie et  $b$  une base de  $E$ , matrice d'un vecteur, d'une famille de vecteurs dans la base  $b$ . Si  $E'$  est un  $\mathbb{K}$ -ev de dimension finie,  $b'$  une base de  $E'$  et  $f \in \mathcal{L}(E, E')$ , matrice de  $f$  dans la base  $b$  au départ et  $b'$  à l'arrivée, notation  $\mathcal{M}_{b',b}(f)$ . Matrice dans  $b$  et  $b'$  de  $f+g$ , de  $\alpha f$ . Matrice associée à  $\text{gof}$  si  $f \in \mathcal{L}(E, E')$  et  $g \in \mathcal{L}(E', E'')$ . Expression matricielle de  $y = f(x)$ .  
Théorème fondamental :  $f \rightarrow \mathcal{M}_{b',b}(f)$  est un isomorphisme, conséquences.  
Si  $A = \mathcal{M}_{b',b}(f)$  avec  $\dim(E) = \dim(E')$  alors  $f$  est un isomorphisme ssi  $A$  est bijective et  $A^{-1} = \mathcal{M}_{b,b'}(f^{-1})$ .  
Cas des endomorphismes et des matrices carrées. Si  $A = \mathcal{M}_{b,b}(f)$  alors  $\text{rg}(A) = \text{rg}(f)$ .
- Application linéaire canoniquement associée à une matrice rectangulaire, endomorphisme canoniquement associée à une matrice carrée. Noyau et Image d'une matrice.  
Une matrice  $A \in \mathcal{M}_n(\mathbb{K})$  est inversible ssi  $\text{Ker}(A) = \{0\}$  ssi  $\text{rg}(A) = n$  ssi ses vecteurs lignes forment une base de  $\mathbb{K}^n$  ssi ses vecteurs colonnes forment une base de  $\mathbb{K}^n$ .  
Méthode d'inversion d'une matrice par résolution du système  $AX = Y$ .  
Inversibilité et inverse d'une matrice diagonale, d'une matrice triangulaire.
- Matrices de passage d'une base  $b$  vers une base  $b'$ , inversibilité et inverse.  
Méthode d'inversion d'une matrice par utilisation des matrices de passage.  
Formules de changement de bases : pour un vecteur, une application linéaire, un endomorphisme.

**Démonstrations au programme de colle :**

- (\*) Sous réserve de compatibilité des formats :  $A \times (B \times C) = (A \times B) \times C$  et  $A \times (B + C) = A \times B + A \times C$ .
- (\*) L'ensemble des matrices triangulaires supérieures est un sous anneau de  $\mathcal{M}_n(\mathbb{K})$  et expression des coefficients diagonaux du produit, transposée d'un produit et transposée d'une matrice inversible
- (\*) Une matrice carrée est inversible ssi elle est inversible à gauche ssi elle est inversible à droite et formules de changement de bases.
- (\*)  $\mathcal{S}_n(\mathbb{K})$  et  $\mathcal{A}_n(\mathbb{K})$  sont des sev de  $\mathcal{M}_n(\mathbb{K})$  et leur dimension,  $\mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K}) = \mathcal{M}_n(\mathbb{K})$ .
- (\*) Matrice associée à  $\text{gof}$  si  $f \in \mathcal{L}(E, E')$  et  $g \in \mathcal{L}(E', E'')$  et si  $f \in \mathcal{L}(E, E')$ ,  $x \in E$  et  $y \in E'$  expression matricielle de  $y = f(x)$ .

FIGURE 14 – Programme de colle n°28

Classe de MPSI 2  
Mathématiques  
Année 2020-2021.  
Semaine du 31/05 au 05/06.

Programme de colle n°29

Matrices :

- Matrices à p lignes et n colonnes à coefficients dans  $\mathbb{K}=\mathbb{R}$  ou  $\mathbb{C}$ . Vocabulaire matriciel.
- L'espace vectoriel  $\mathcal{M}_{p,n}(\mathbb{K})$ , base canonique, dimension, sous espaces des matrices diagonales, triangulaires supérieures (resp inférieures).
- Produit matriciel : définition et propriétés.
- L'anneau  $\mathcal{M}_n(\mathbb{K})$ , non commutativité pour  $n \geq 2$ , puissances de matrices, matrices nilpotentes, matrices qui commutent, formule du binôme et factorisation  $A^*B^*$  pour deux matrices carrées qui commutent. Le sous-anneau des matrices diagonales et des matrices triangulaires supérieures (resp inférieures).
- Matrices carrées inversibles : définition,  $I_n$  est inversible, si A est inversible alors  $A^{-1}$  est inversible et  $(A^{-1})^{-1}=A$ , si A et B sont inversibles alors  $A \cdot B$  est inversible et expression de  $(A \cdot B)^{-1}$ .
- Une matrice inversible est régulière à gauche et à droite.
- Une matrice carrée est inversible ssi elle est inversible à gauche ssi elle est inversible à droite.
- Rang d'une matrice : c'est le rang de ses vecteurs colonnes. Si  $A \in \mathcal{M}_{p,n}(\mathbb{K})$  alors  $\text{rg}(A) \leq \min(p, n)$ ,  $\text{rg}(A)$  ssi les n colonnes de A forment une famille libre de  $\mathbb{K}^p$  et  $\text{rg}(A) = p$  ssi les n colonnes de A forment une famille génératrice de  $\mathbb{K}^p$ .
- Transposée d'une matrice : Définition, propriétés, linéarité, transposée d'un produit, transposée d'une matrice inversible. La transposition est un isomorphisme de  $\mathcal{M}_{p,n}(\mathbb{K})$  dans  $\mathcal{M}_{n,p}(\mathbb{K})$  et un automorphisme de  $\mathcal{M}_n(\mathbb{K})$ . Matrices symétriques et antisymétriques : définitions, notation,  $s_n(\mathbb{K})$  et  $a_n(\mathbb{K})$  sont des sev supplémentaires de  $\mathcal{M}_n(\mathbb{K})$ , dimension de ces sous-espaces.
- Si E est un  $\mathbb{K}$ -ev de dimension finie et b une base de E, matrice d'un vecteur, d'une famille de vecteurs dans la base b, Si E' est un  $\mathbb{K}$ -ev de dimension finie, b' une base de E', et  $f \in \mathcal{L}(E, E')$ , matrice de f dans la base b au départ et b' à l'arrivée, notation  $\mathcal{M}_{b,b'}(f)$ . Matrice dans b et b' de  $\text{Fg}$ , de  $\text{dF}$ . Matrice associée à  $\text{pod}$  si  $f \in \mathcal{L}(E, E')$  et  $g \in \mathcal{L}(E', E)$ . Expression matricielle de  $y=f(x)$ .
- Théorème fondamental :  $f: \mathcal{M}_{b,b'}(f)$  est un isomorphisme, conséquences.
- Si  $A = \mathcal{M}_{b,b'}(f)$  avec  $\dim(E) = \dim(E')$  alors f est un isomorphisme ssi A est bijective et  $A^{-1} = \mathcal{M}_{b',b}(f^{-1})$ . Cas des endomorphismes et des matrices carrées. Si  $A = \mathcal{M}_{b,b}(f)$  alors  $\text{rg}(A) = \text{rg}(f)$ .
- Application linéaire canoniquement associée à une matrice rectangulaire, endomorphisme canoniquement associée à une matrice carrée. Noyau et Image d'une matrice.
- Une matrice  $A \in \mathcal{M}_n(\mathbb{K})$  est inversible ssi  $\text{Ker}(A) = \{0\}$  ssi  $\text{rg}(A) = n$  ssi ses vecteurs lignes forment une base de  $\mathbb{K}^n$ .
- Méthode d'inversion d'une matrice par résolution du système  $AX=Y$ .
- Inversibilité et inverse d'une matrice diagonale, d'une matrice triangulaire.
- Matrices de passage d'une base b vers une base b', Inversibilité et inverse.
- Méthode d'inversion d'une matrice par utilisation des matrices de passage.
- Formules de changement de bases : pour un vecteur, une application linéaire, un endomorphisme.
- Matrices blocs, interprétation en termes de stabilité, règle de calculs par blocs.
- Matrices équivalentes, c'est une relation d'équivalence, caractérisation à l'aide des applications linéaires : deux matrices sont équivalentes ssi elles représentent la même application linéaire dans deux couples de bases différentes.  $A \in \mathcal{M}_{p,n}(\mathbb{K})$  :  $\text{rg}(A) = r$  ssi A est équivalente à la matrice

$$J = \begin{pmatrix} 1 & & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & 0 \\ & & & & \\ & & & & \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix} \text{ } J = \begin{pmatrix} 1 & & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & 0 \\ & & & & \\ & & & & \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix} \text{ } J = \begin{pmatrix} 1 & & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & 0 \\ & & & & \\ & & & & \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix} \text{ } J = \begin{pmatrix} 1 & & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & 0 \\ & & & & \\ & & & & \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix}$$

est donc aussi le rang de ses vecteurs lignes. Le rang d'une matrice extraite de A est inférieur ou égal au rang de A.  $A \in \mathcal{M}_{p,n}(\mathbb{K})$  est de rang r ssi il existe une matrice  $A' \in \mathcal{M}_r(\mathbb{K})$  extraite de A inversible et toute matrice  $A'' \in \mathcal{M}_r(\mathbb{K})$  avec  $s \cdot r \cdot n$  est pas inversible.

• Matrices semblables, c'est une relation d'équivalence, deux matrices sont semblables ssi elles représentent le même endomorphisme dans deux bases différentes.

• Trace d'une matrice carrée : définition, propriétés, Trace d'un endomorphisme : définition, propriétés.

• Opérations élémentaires sur colonnes (ou lignes) d'une matrice : permutation de deux colonnes, multiplication d'une colonne par un scalaire non nul, addition d'une colonne à une autre colonne.

• Matrices associées à des opérations élémentaires : effectuer une opération élémentaire sur les colonnes (resp lignes) de A c'est multiplier A à droite (resp à gauche) par la matrice correspondante.

Les matrices associées aux opérations élémentaires sont inversibles. Les opérations élémentaires appliquées aux colonnes de A conservent  $\text{Im}(A)$ , appliquées aux lignes de A elles conservent  $\text{Ker}(A)$ . Les opérations élémentaires appliquées aux lignes ou colonnes de A conservent le rang de A.

• Détermination pratique du rang d'une matrice par la méthode du pivot de Gauss.

• Cas particulier des matrices carrées inversibles : détermination pratique de l'inverse par la méthode du pivot de Gauss sur les colonnes de A uniquement (ou les lignes de A uniquement).

• Retour sur les systèmes linéaires : écriture matricielle, écriture vectorielle.

• Existence et unicité des solutions d'un système en fonction de son rang.

Démonstrations au programme de colle :

- (\*) Si  $A \in \mathcal{M}_n(\mathbb{K})$  pour  $A = \begin{pmatrix} a & & 0 \\ & b & \\ c & c' & c'' \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$  avec :  $a^{-1}b^{-1}c'' = a^{-1}b^{-1}c'' = a^{-1}b^{-1}c'' = 1$  et  $aa^{-1}bb^{-1}cc'' = aa^{-1}bb^{-1}cc'' = aa^{-1}bb^{-1}cc'' = 0$ , montrer que A est inversible et trouver l'inverse de A.
- (\*)  $A \in \mathcal{M}_{p,n}(\mathbb{K})$  :  $\text{rg}(A) = r$  ssi A est équivalente à la matrice  $J = \begin{pmatrix} 1 & & 0 & \dots & 0 \\ 0 & \dots & 0 & \dots & 0 \\ & & & & \\ & & & & \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix}$ .
- (\*)  $A \in \mathcal{M}_{p,n}(\mathbb{K})$  est de rang r ssi  $(\exists A' \in \text{GL}_r(\mathbb{K}))$  extraite de A  $(\forall A'' \in \mathcal{M}_r(\mathbb{K}))$  avec  $s \cdot r \cdot n$  extraite de A,  $A'' \in \text{GL}_r(\mathbb{K})$ .
- (\*) Propriétés de la trace d'une matrice : linéarité,  $\text{Tr}(AB) = \text{Tr}(BA)$  et 2 matrices semblables ont même trace.
- (\*) La trace d'un projecteur est son rang et si  $f, g \in \mathcal{L}(E)$  tels que  $\text{log} = \text{rg}$  alors  $\text{Ker}(f)$  et  $\text{Im}(f)$  sont stables par g.

FIGURE 15 – Programme de colle n°29

Classe de MPSI 2  
Mathématiques

Année 2020-2021.  
Semaine du 07/06 au 12/06.

**Programme de colle n°30**

**Variables aléatoires sur un univers fini :**

- Définition d'une variable aléatoire (VA) sur un univers fini  $\Omega$ , variable aléatoire réelle (VAR).  
Exemples : variable aléatoire constante, variable aléatoire indicatrice.  
Notations :  $(X=a)$  et si  $X$  est à valeurs réelles  $(X < a)$ ,  $(X > a)$ ,  $(X \leq a)$ ,  $(X \geq a)$ .  
Si  $X$  est une v.a et  $f$  une application définie sur  $X(\Omega)$  alors  $f \circ X$  est une v.a.
- Loi d'une variable aléatoire.  
Si  $\Omega$  est un univers fini de cardinal  $m \geq n$ ,  $E = \{x_1, \dots, x_n\}$  et  $(p_k)_{1 \leq k \leq n}$  sont des réels positifs de somme 1 alors il existe  $X$  une v.a sur  $\Omega$  et  $P$  une probabilité sur  $\Omega$  tel que  $X(\Omega) = \{x_k, 1 \leq k \leq n\}$  et  $\forall k \in [1, n]$   
 $P(X = x_k) = p_k$ . Si  $X$  est une v.a à valeurs dans  $\mathbb{Z}$  alors  $\forall k \in \mathbb{Z} : P(X = k) = P(X \leq k) - P(X \leq k - 1) = P(X \geq k) - P(X \geq k + 1)$ .
- Lois usuelles : loi uniforme, loi de Bernoulli et loi binomiale : Interprétation et exemples.
- Couple de variables aléatoires noté  $(X, Y)$ . Loi conjointe du couple et lois marginales du couple  $(X, Y)$ .  
La loi conjointe détermine entièrement les lois marginales mais la réciproque est fausse.  
Loi conditionnelle de  $X$  sachant  $(Y = y)$ . Généralisation aux  $n$ -uplets de v.a :  $(X_1, \dots, X_n)$   
Indépendance de  $X$  et  $Y$ , caractérisation par les lois conditionnelles sachant  $(Y = y)$  ou  $(X = x)$ .  
Si  $X$  et  $Y$  sont indépendantes et  $f$  une fonction définie sur  $X(\Omega)$  et  $g$  une fonction définie sur  $Y(\Omega)$  alors  $f(X)$  et  $g(Y)$  sont indépendantes. Suites de v.a mutuellement indépendantes. Lemme des coalitions.  
Interprétation de la loi binomiale comme sommes de  $n$  v.a mutuellement indépendantes suivant la même loi de Bernoulli.
- Espérance d'une VAR sur un univers. Espérance des lois usuelles.  
Théorème de transfert. Linéarité, positivité et croissance de l'espérance.  
Si  $X$  et  $Y$  sont indépendantes alors  $E(XY) = E(X)E(Y)$  réciproque fausse.
- Moment d'ordre  $r$  d'une v.a, variance d'une v.a. Formule de Koenig-Huygens.  
Variance des lois usuelles. Variable aléatoire centrée réduite.  $V(aX + b) = a^2 V(X)$ .
- Covariance d'un couple de v.a et coefficient de corrélation. Formule de Koenig-Huygens, si  $X$  et  $Y$  sont indépendantes alors  $\text{Cov}(X, Y) = 0$  et la réciproque est fausse.  
Symétrie et bilinéarité de la covariance :  $\text{Cov}(X, X) = V(X)$ .  
 $V(X + Y)$  cas général et cas où  $X$  et  $Y$  sont indépendantes,  $V(aX + bY)$ .  
Généralisation à  $n$  variables aléatoires. Inégalité de Cauchy Schwarz et  $-1 \leq \rho(X, Y) \leq 1$ .
- Inégalité de Markov et Inégalité de Bienaymé Tchebychev.

Démonstrations au programme de colle.

- (\*) Sommes de  $n$  v.a mutuellement indépendantes suivant la loi binomiale  $m_i, p$ .
- (\*) Si  $X$  est une VAR alors  $E(X) = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\})$ , théorème de transfert, linéarité, positivité et croissance de l'espérance.
- (\*) Espérance et variance des lois usuelles (2 méthodes pour la loi binomiale dont une est la somme de  $n$  VA mutuellement indépendantes de Bernoulli).
- (\*) Si  $X$  et  $Y$  sont indépendantes alors  $E(XY) = E(X)E(Y)$ ,  $V(aX + bY)$  et  $V(X_1 + \dots + X_n)$ .
- (\*) Inégalités : Cauchy-Schwarz, Markov et Bienaymé Tchebychev.

FIGURE 16 – Programme de colle n°30



## B. Algorithme DQN - Implémentation du jeu

Tout les programmes et codes présentés ici sont accessibles sur [https://github.com/aurianegbt/Reinforcment\\_Learning\\_DQN](https://github.com/aurianegbt/Reinforcment_Learning_DQN). Dans l'annexe qui suit, la construction de l'environnement est commentée.

On propose de tester les algorithmes et différentes versions sur un jeu plus simple : le *Snake*. Le principe du jeu est le suivant : un serpent (le snake) se déplace sur une grille à la recherche des pommes, il doit en manger le plus possible, mais sans toucher ni les murs, ni son propre corps, sans quoi il meurt. A chaque pomme mangée, il grandit de 1, augmentant la difficulté du jeu. Le serpent peut se déplacer sur toute la grille en allant tout droit, à sa gauche ou à sa droite par rapport à sa direction. On implémente le jeu en définissant une classe `Snake_environment` et les méthodes suivantes :

a. Le jeu est initialisé par des paramètres initiaux (taille de la grille, taille initiale du serpent, gain, pénalité...) correspondant à la fonction `__init__` de notre *Snake*. Une fois les constantes du jeu décidées, on définit la grille de jeux et positionne le serpent (le plus au milieu possible et dirigé vers le haut) sur celle-ci, on positionne une première pomme sur l'écran. On décide comme convention que la carte sera notée d'un 0 si la case est libre, d'un 1 si elle est occupée par une pomme, et 0.5 si c'est par le serpent. Pour connaître la direction dans laquelle se dirige le serpent, on garde en mémoire le dernier mouvement : en effet, si le serpent vient d'avancer vers une direction, c'est qu'il est dirigé à présent dans cette direction. La position du serpent est implémentée par une liste dont la longueur correspond à la taille du serpent et la tête est située au début de la liste.

```
[1]: class Snake_environment:
    def __init__(self):
        self.row = 8                # Nombre de lignes
        self.col = 8                # Nombre de colonnes
        self.snake_initial = 2
        self.time_penalty = -3      # Penalité de temps
        self.death_penalty = -100   # Pénalité de mort
        self.apple_reward = 200     # Gain

        self.snake_position = []

        self.map = np.zeros((self.row, self.col))

        for i in range(self.snake_initial):
            row_init = int(self.row/2)+i
            col_init = int(self.col/2)
            self.snake_position+=(row_init,col_init)]
            self.map[row_init,col_init] = 0.5
```

```

self.apple_position = self.new_apple()

self.collected = False
self.last_action = 0

```

b. Le placement des pommes est aléatoire sur la grille et sera implémenté par la fonction `new_apple`.

```

[2]: def new_apple(self):
    new_row = np.random.randint(0,self.row)
    new_col = np.random.randint(0,self.col)

    while self.map[new_row,new_col]==0.5 : # i.e. le serpent est dessus
        new_row = np.random.randint(0,self.row)
        new_col = np.random.randint(0,self.col)

    self.map[new_row,new_col] = 1
    return(new_row,new_col)

```

c. Lorsque le serpent se déplace, la méthode `snake_move` met à jour la position du serpent en fonction de la prochaine position. Si le serpent a attrapé une pomme, alors celui-ci grandit de 1. au niveau de sa queue. Il faut alors réinitialiser la grille avec la nouvelle position du serpent, et aussi la position de la (nouvelle) pomme.

```

[3]: def snake_move(self,next_row,next_col,apple_collected):
    self.snake_position.insert(0,(next_row,next_col))

    if not apple_collected :
        self.snake_position.pop()

    elif apple_collected :
        self.collected = True

    self.map = np.zeros((self.row,self.col))

    for i in range(len(self.snake_position)):
        pos_row = self.snake_position[i][0]
        pos_col = self.snake_position[i][1]
        self.map[pos_row,pos_col]=0.5

    apple_pos_row = self.apple_position[0]
    apple_pos_col = self.apple_position[1]
    self.map[apple_pos_row,apple_pos_col]= 1

```

d. Maintenant, le principe du jeu est encodé, mais il est nécessaire de faire une fonction qui permet au jeu d'avancer. Cette fonction `step` permet de faire avancer le jeu d'une étape en fonction de l'action fournie par l'agent. Selon l'action effectuée, la fonction permet de déplacer le serpent, calculer les gains, vérifier que le serpent n'est pas dans une situation de Game Over. On prend la convention suivante :

$\mathcal{A} = \{0, 1, 2, 3\}$  où 0 correspond au déplacement vers le haut, 1 celui vers le bas, 2 vers la droite et enfin 3 vers la gauche. On remarque alors quelques contraintes de déplacement selon la position précédente du serpent. En effet, toutes les actions ne sont pas réalisables dans n'importe quel ordre : par exemple, le serpent ne peut pas faire un demi-tour complet, donc s'il est allé dans une direction à un instant, il ne pourra pas aller dans la direction opposée à l'instant suivant. On décide dans ces cas là que l'action décidée, irréalisable, est remplacée par défaut par la précédente. Donc, si par exemple, l'agent souhaite aller à gauche après être allé à droite, l'action décidée est remplacée par aller à droite.

```
[1]: def step(self, action):
    #de base, on réinitialise ça : (en premier si jamais il rencontre une pomme,
    #ensuite)
    self.collected = False

    if action == 3 and self.last_action == 2 :
        action = 2
    elif action == 0 and self.last_action == 1 :
        action = 1
    elif action == 2 and self.last_action == 3 :
        action = 3
    elif action == 1 and self.last_action == 0 :
        action = 0

    pos_row = self.snake_position[0][0]
    pos_col = self.snake_position[0][1]

    if action == 0 :
        next_row = pos_row-1
        next_col = pos_col
        #Le serpent perd s'il rencontre un mur ou lui-même:
        #Grille de jeu : 0 = rien, 0.5 = serpent, 1 = pomme
        if next_row < 0 :
            game_over = True
            reward = self.death_penalty
        else:
            if self.map[next_row,next_col] == 0.5:
                game_over = True
                reward = self.death_penalty
            if self.map[next_row,next_col] == 1:
                game_over = False
                reward = self.apple_reward
                self.snake_move(next_row,next_col, apple_collected=True)
            if self.map[next_row,next_col] == 0:
                game_over = False
                reward = self.time_penalty
                self.snake_move(next_row,next_col, apple_collected = False)

    elif action == 1 :
        :
```

```
self.last_action = action

return(self.map,reward,game_over)
```

e. Enfin, pour terminer, il faut une fonction qui réinitialise le jeu, pour chaque début de partie. Celle-ci est implémentée dans `reset`.

```
[4]: def reset(self):
      self.map = np.zeros((self.row,self.col))
      self.snake_position = []

      for i in range(self.snake_initial):
          row_init = int(self.row/2)+i
          col_init = int(self.col/2)

          self.snake_position+=(row_init,col_init)]

          self.map[row_init,col_init] = 0.5

      apple_pos_row = self.apple_position[0]
      apple_pos_col = self.apple_position[1]
      self.map[apple_pos_row,apple_pos_col]= 1

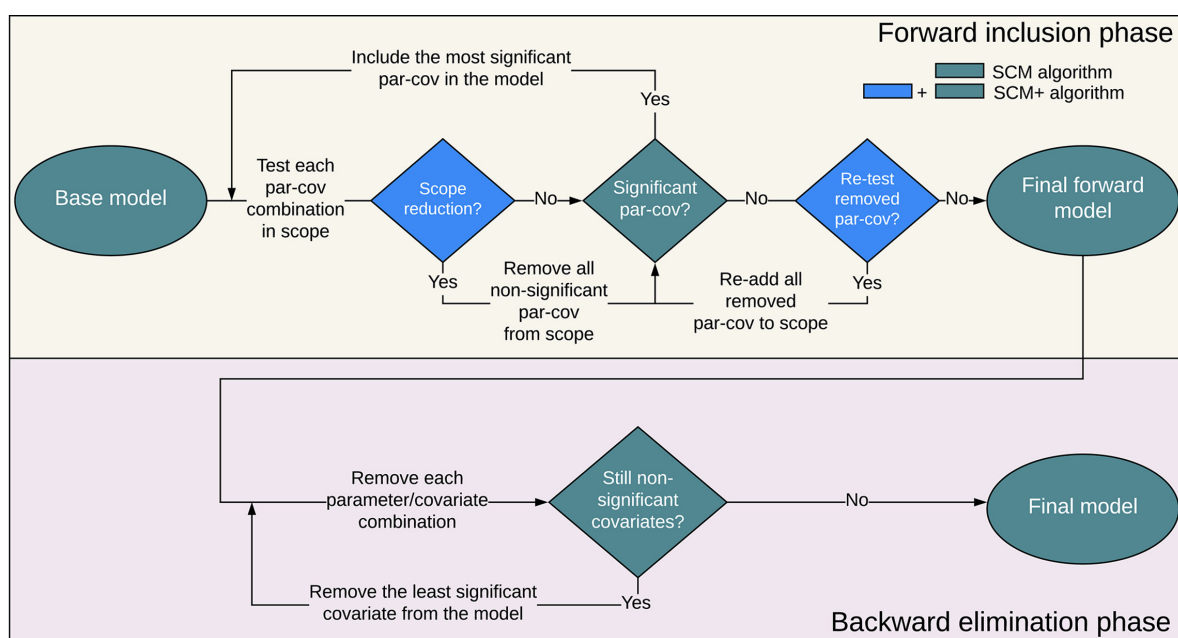
      self.last_move = 0
```



## C. Algorithme de construction de Modèle

### C.1. Détails sur les algorithmes de sélection de covariables

SCM :



CPT Pharmacom & Syst Pharma, Volume: 11, Issue: 9, Pages: 1210-1222, First published: 19 July 2022, DOI: (10.1002/psp4.12838)

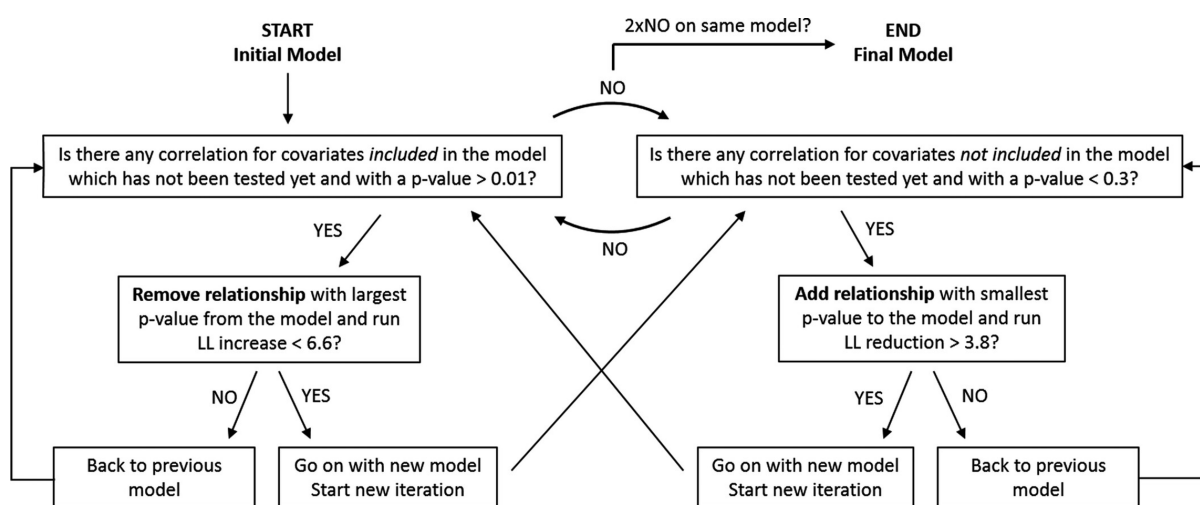
FIGURE 17 – Vue d'ensemble des algorithmes Stepwise Covariate Modelling (SCM) et Stepwise Covariate Modelling amélioré (SCM+), extrait de l'article [6].

Le principe de base de l'algorithme SCM, décrit en figure 17, est une approche itérative de sélection de covariables. L'algorithme se compose de deux étapes : l'ajout de covariables *-forward-* et l'élimination de covariables *-backward-*. Initialement, le modèle ne contient aucune covariable. Lors de l'étape d'ajout de covariables *-forward-*, chaque covariable est testée séparément pour voir si son inclusion améliore significativement la vraisemblance du modèle. Celles l'améliorant sont conservées dans le modèle. Le processus se répète alors avec l'ajout séquentiel d'autres covariables jusqu'à ce que l'inclusion de nouvelles covariables ne donne plus d'amélioration significative. Ensuite, l'étape d'élimination de covariables *-backward-* commence. Le processus consiste maintenant à éliminer séquentiellement chaque covariable du modèle pour évaluer si son retrait améliore le modèle ou non. Si l'élimination d'une covariable n'entraîne pas de

détérioration significative, cette covariable est retirée du modèle. Le processus se poursuit jusqu'à ce que le retrait de covariables n'améliore plus de manière significative le critère d'ajustement.

L'algorithme SCM+ présente lui plusieurs améliorations par rapport au SCM. Tout d'abord, il intègre une adaptation de la réduction du champ d'application, permettant ainsi d'éliminer certaines covariables dès les premières étapes si elles ne montrent pas de bon résultat pour améliorer le modèle. En limitant le nombre d'évaluations de fonctions, le SCM+ évite des itérations excessives sans amélioration significative du modèle, ce qui peut se produire dans le SCM.

## COSSAC



**CPT Pharmacom & Syst Pharma, Volume: 10, Issue: 4, Pages: 318-329, First published: 23 March 2021, DOI: (10.1002/psp4.12612)**

FIGURE 18 – Schéma de la procédure de construction de modèle de covariables COnditional Sampling use for Stepwise Approach based on Correlation tests (COSSAC). La notation LL représente la log-vraisemblance multiplié par  $-2$ . Figure extraite de [2]

Le principe de l'algorithme COSSAC, décrit en figure 18 est de sélectionner de manière itérative les covariables les plus pertinentes en utilisant des tests statistiques basés sur des échantillons des distributions conditionnelles des paramètres individuels. Contrairement au SCM, COSSAC ne nécessite pas de réestimer le modèle complet à chaque étape d'ajout ou de suppression de covariables, ce qui permet de réduire considérablement le temps de calcul.

Le processus COSSAC comprend deux étapes principales : l'étape d'ajout de covariables *-forward-* et l'étape d'élimination de covariables *-backward-*. À chaque itération, l'algorithme COSSAC teste séquentiellement les covariables pour déterminer leur corrélation avec les paramètres du modèle. Les covariables ayant des corrélations significatives sont ajoutées au modèle. Une fois que des covariables ont été incluses dans le modèle, l'algorithme COSSAC vérifie séquentiellement si certaines covariables peuvent être supprimées sans détériorer significativement la qualité du modèle. Les covariables dont la suppression n'affecte pas de manière significative le critère d'ajustement sont retirées du modèle.

Les avantages de l'algorithme COSSAC sont les suivants. L'algorithme COSSAC nécessite beaucoup moins d'itérations et de réestimations du modèle complet par rapport à des approches plus exhaustives

comme l'algorithme SCM, ce qui permet de réduire considérablement le temps de calcul, en particulier pour les modèles complexes. L'utilisation de tests statistiques basés sur des échantillons des distributions conditionnelles permet de contrôler le risque de sur-ajustement et d'éviter de sélectionner des covariables qui ne sont pas réellement significatives.

## C.2. Algorithme de validation croisée $K$ -folds

Maintenant, pour estimer efficacement l'erreur de prédiction pour un paramètre  $\lambda_l$ ,  $l \leq L$ , une méthode de validation croisée est utilisée. Elle consiste à diviser l'ensemble de données en plusieurs parties (ou "folds") pour estimer l'efficacité du modèle de manière plus fiable. À chaque itération, l'un des folds est exclu de l'ensemble d'entraînement, c'est-à-dire qu'il n'est pas utilisé pour ajuster les paramètres du modèle aux données. Ensuite, ce fold exclu est utilisé comme ensemble de validation, c'est-à-dire que les données de ce fold sont utilisées pour évaluer les performances du modèle en tant que prédicteur. Ce processus est répété pour chaque folds, en utilisant chaque fold une fois comme ensemble de validation, pendant que les autres folds sont utilisés pour l'entraînement du modèle. Finalement, les performances obtenues à chaque itération sont agrégées pour donner une estimation plus fiable de la capacité du modèle à généraliser à de nouvelles données et à éviter le surajustement. L'algorithme pour  $K$  folds est donnée en table 9.

---

### Algorithme 9 : $K$ -folds Cross-Validation [3]

---

**Data :** Soit un modèle  $\mathcal{M}$ , des données  $(X_i, Y_i)_{i \leq N}$

L'ensemble  $\mathcal{I} = \{1, \dots, N\}$  est partitionner aléatoirement en  $K$  sous-ensemble  $(\mathcal{I}_k)_{k \leq K}$  de cardinal le plus égal possible ;

**pour**  $k \leq K$  **faire**

    Les paramètres du modèle  $\mathcal{M}$  sont ajustés en utilisant les données  $(X_i, Y_i)_{i \in \mathcal{I} \setminus \mathcal{I}_k}$  ;

    L'erreur de prédiction PE du modèle  $\mathcal{M}$  est estimé en utilisant l'erreur quadratique moyenne MSE à partir des données  $(X_i, Y_i)_{i \in \mathcal{I}_k}$ , cette dernière est noté  $MSE_k$  ;

**fin**

L'erreur quadratique moyenne pour le modèle  $\mathcal{M}$  est alors donnée par

$$\overline{MSE} = \frac{1}{K} \sum_{k \leq K} MSE_k \quad (1)$$


---

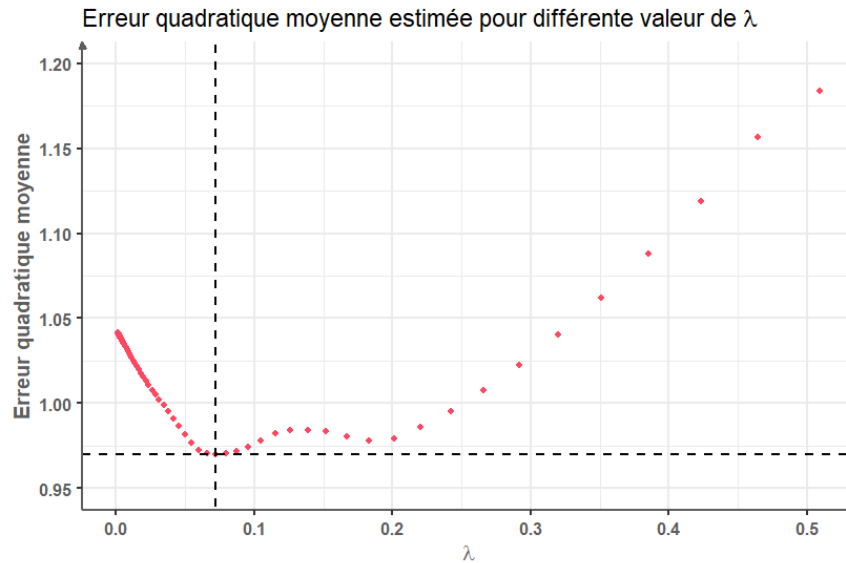


FIGURE 19 – Estimation de l’erreur quadratique moyenne pour différente valeur de  $\lambda$  pour des données simulées du modèle (MOD AB).

L’algorithme 9 est finalement appliqué à chacun des  $\lambda_l$ ,  $l \leq L$  et nous obtenons alors une estimation de l’erreur quadratique moyenne (1) pour chaque  $\lambda_l$ , comme représenté en figure 19. Le paramètre de pénalisation  $\lambda$  choisit est alors égal au  $\lambda_l$  minimisant l’erreur de prédiction.

### C.3. Paramètres individuels

| Paramètre                                  | Définition                                              | Valeur moyenne       |
|--------------------------------------------|---------------------------------------------------------|----------------------|
| <b>Effets Fixes</b>                        |                                                         |                      |
| $\delta_{Ab}$                              | taux de décroissance des anticorps (par jour)           |                      |
| Chez les femmes                            |                                                         | 0.0251               |
| Chez les hommes                            |                                                         | 0.0353               |
| $\delta_S$                                 | taux de décroissance des cellules $S$ (par jour)        | 0.231                |
| $\delta_L$                                 | taux de décroissance des cellules $L$ (par jour)        | $3.16 \cdot 10^{-4}$ |
| $\varphi_S$                                | afflux des cellules $S$ (Eu/mL/jour)                    |                      |
| A l’âge moyen 35 ans                       |                                                         | 3057                 |
| $FC\Delta AGE = +1an$                      | facteur multiplicatif pour une différence d’âge de 1 an | 0.934                |
| $\varphi_L$                                | afflux des cellules $L$ (Eu/mL/jour)                    |                      |
| Chez les participants d’origine africaine  |                                                         | 16.6                 |
| Chez les participants d’origine européenne |                                                         | 70.7                 |
| <b>Effets aléatoires</b>                   |                                                         |                      |
| $\omega_S$                                 | écart-type des effets aléatoires sur $\varphi_S$        | 0.92                 |
| $\omega_L$                                 | écart-type des effets aléatoires sur $\varphi_L$        | 0.85                 |
| $\omega_{Ab}$                              | écart-type des effets aléatoires sur $\delta_{Ab}$      | 0.30                 |
| <b>Modèle d’erreur</b>                     |                                                         |                      |
| $\sigma_{Ab}$                              | écart-type du modèle d’erreur                           | 0.10                 |

TABLE 1 – Valeur des paramètres  $\psi$  utilisée pour les simulations, d’après [1].

## C.4. Génération des covariables individuels

Pour générer les covariables supplémentaires, une fonction `randomCovariate` fournissant une distribution aléatoire. Cette fonction donne en sortie une liste contenant le nom de la distribution et les éléments nécessaire à son échantillonnage. Si par exemple la distribution est “unif” pour une loi uniforme les éléments seront “min” et “max” pour les bornes de l’intervalle. Le tableau 2 résume le procédé de générations de chacune des lois. Pour donc générer une distribution pour une variable aléatoire  $X$ , alors l’appel à la fonction permet de tirer uniformément parmi les lois possibles, citées dans le tableau 2, puis tire les paramètres ensuite selon les lois renseignées. Les covariables de chaque individus sont ensuite tirés selon cette loi.

| Suffixe pour la loi | Élément(s) | Génération                                                                                                                                                |
|---------------------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|
| Loi beta            |            |                                                                                                                                                           |
| “beta”              | “shape1”   | $\mathcal{P}(20)$                                                                                                                                         |
|                     | “shape2”   | $\mathcal{P}(20)$                                                                                                                                         |
| Loi binomiale       |            |                                                                                                                                                           |
| “binom”             | “prob”     | $\mathcal{U}([0.2; 0.8])$                                                                                                                                 |
|                     | “size”     | $(\mathcal{P}(10) \vee 20) \wedge 1$                                                                                                                      |
| Loi de Cauchy       |            |                                                                                                                                                           |
| “cauchy”            | “location” | $\mathcal{N}(0, 20^2)$                                                                                                                                    |
|                     | “scale”    | $\mathcal{E}(0.1)$                                                                                                                                        |
| Loi du $\chi^2$     |            |                                                                                                                                                           |
| “chisq”             | “df”       | $\mathcal{P}(20) \wedge 3$                                                                                                                                |
| Loi exponentielle   |            |                                                                                                                                                           |
| “exp”               | “rate”     | $\mathcal{E}(2)$                                                                                                                                          |
| Loi de Fischer      |            |                                                                                                                                                           |
| “f”                 | “df1”      | $\mathcal{P}(20) \wedge 3$                                                                                                                                |
|                     | “df2”      | $\mathcal{P}(20) \wedge 3$                                                                                                                                |
| Loi du gamma        |            |                                                                                                                                                           |
| “gamma”             | “shape”    | $\mathcal{P}(10)$                                                                                                                                         |
|                     | “scale”    | $\mathcal{E}(1/2)$                                                                                                                                        |
| Loi géométrique     |            |                                                                                                                                                           |
| “geom”              | “prob”     | $\mathcal{U}([0; 0.6])$                                                                                                                                   |
| Loi catégorielle    |            |                                                                                                                                                           |
| “sample”            | “size”     | $\mathcal{P}(5) \wedge 2$                                                                                                                                 |
|                     | “prob”     | Par itération : $p_1 \sim \mathcal{U}([0; 1])$ ,<br>$p_k \sim \mathcal{U}([0; 1 - \sum_{i \leq k-1} p_i])$ , et $p_{size} = 1 - \sum_{i \neq size} p_i$ . |
| Loi normale         |            |                                                                                                                                                           |
| “norm”              | “mean”     | $\mathcal{N}(0, 1)$                                                                                                                                       |
|                     | “sd”       | $\mathcal{E}(0.1)$                                                                                                                                        |
| Loi de poisson      |            |                                                                                                                                                           |

| “pois”       | “lambda” | $\mathcal{E}(0.1)$          |
|--------------|----------|-----------------------------|
| Loi uniforme |          |                             |
| “unif”       |          | $x = \mathcal{N}(0_2, I_2)$ |
|              | “min”    | $\min(x)$                   |
|              | “max”    | $\max(x)$                   |

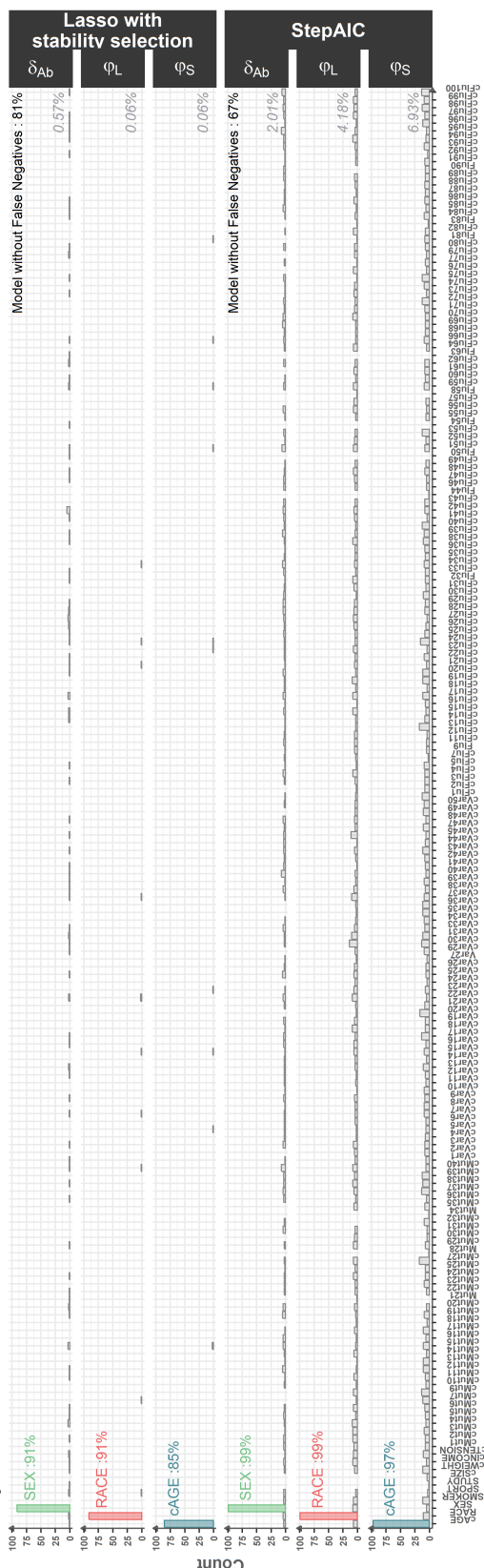
TABLE 2 – Procédé de génération des différentes loi par la fonction `randomCovariate` implémenté

Pour les scénarios avec corrélations, le package `simstudy` [4] a été utilisé. Il permet de générer de nouvelles variables, en liant leur distribution à des covariables préexistantes. Il faut pour cela remplir une table R avec pour chaque élément, les informations sur une variable à générer. La table de définition des données comprend une ligne pour chaque variable à générer, avec les champs suivants : “varname”, “formula”, “variance”, “dist” et “link”. “varname” fournit le nom de la variable à générer. La formule “formula” est soit une valeur, soit une chaîne de caractères représentant n’importe quelle formule R valide (pouvant inclure des appels de fonctions) qui définit dans la plupart des cas la moyenne de la distribution. La variance “variance” est une valeur ou une chaîne de caractères qui spécifie soit la variance, soit un autre paramètre qui caractérise la distribution ; la valeur par défaut est 0. “dist” définit la distribution de la variable à générer ; la valeur par défaut est normale. Le lien “link” est une fonction qui définit la relation de la formule avec la valeur moyenne et peut être soit l’identité, le logarithme ou le logit, selon la distribution ; la valeur par défaut est l’identité. Un exemple d’estimation des taux de corrélations par méthode de Spearman [5] pour les jeux de données de 100 covariables créées est représentées en annexe C.5, sur les figures 25 et 24.



# Covariate Selection Count

With uncorrelated covariates,  
among 100 simulated datasets.



With correlated covariates,  
among 100 simulated datasets.

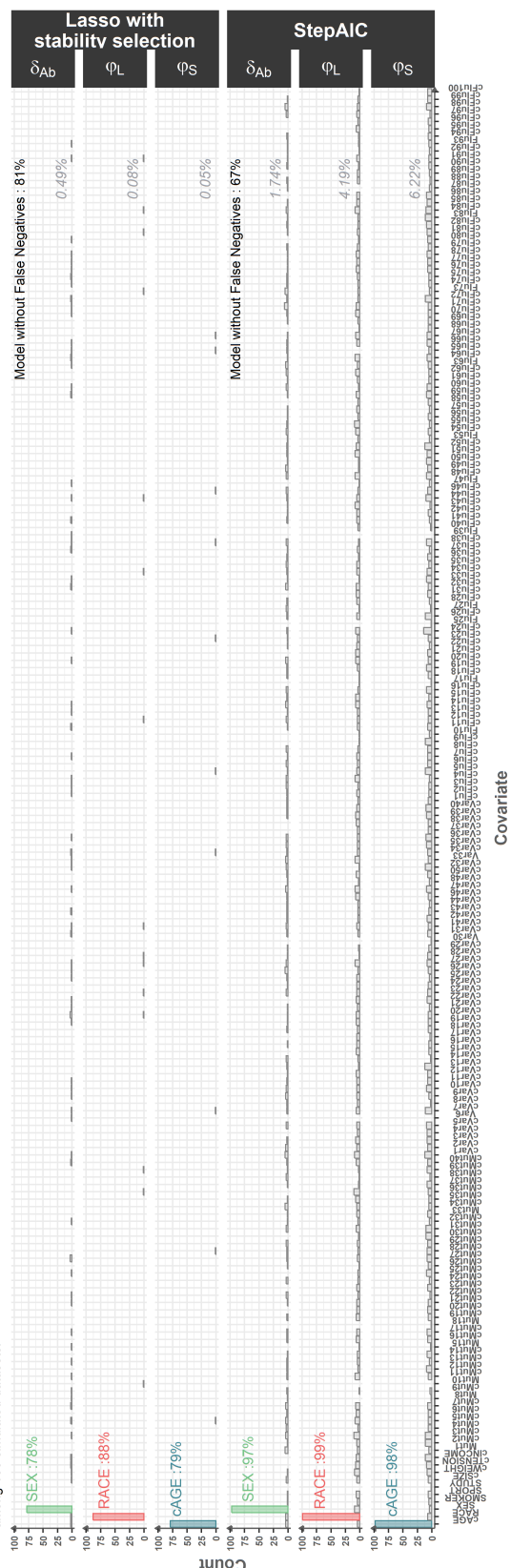


FIGURE 21 – Fréquence de sélection des différentes covariables avec chacun des paramètres, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations. Pour 200 covariables, 100 individus.



# Covariate Selection Count

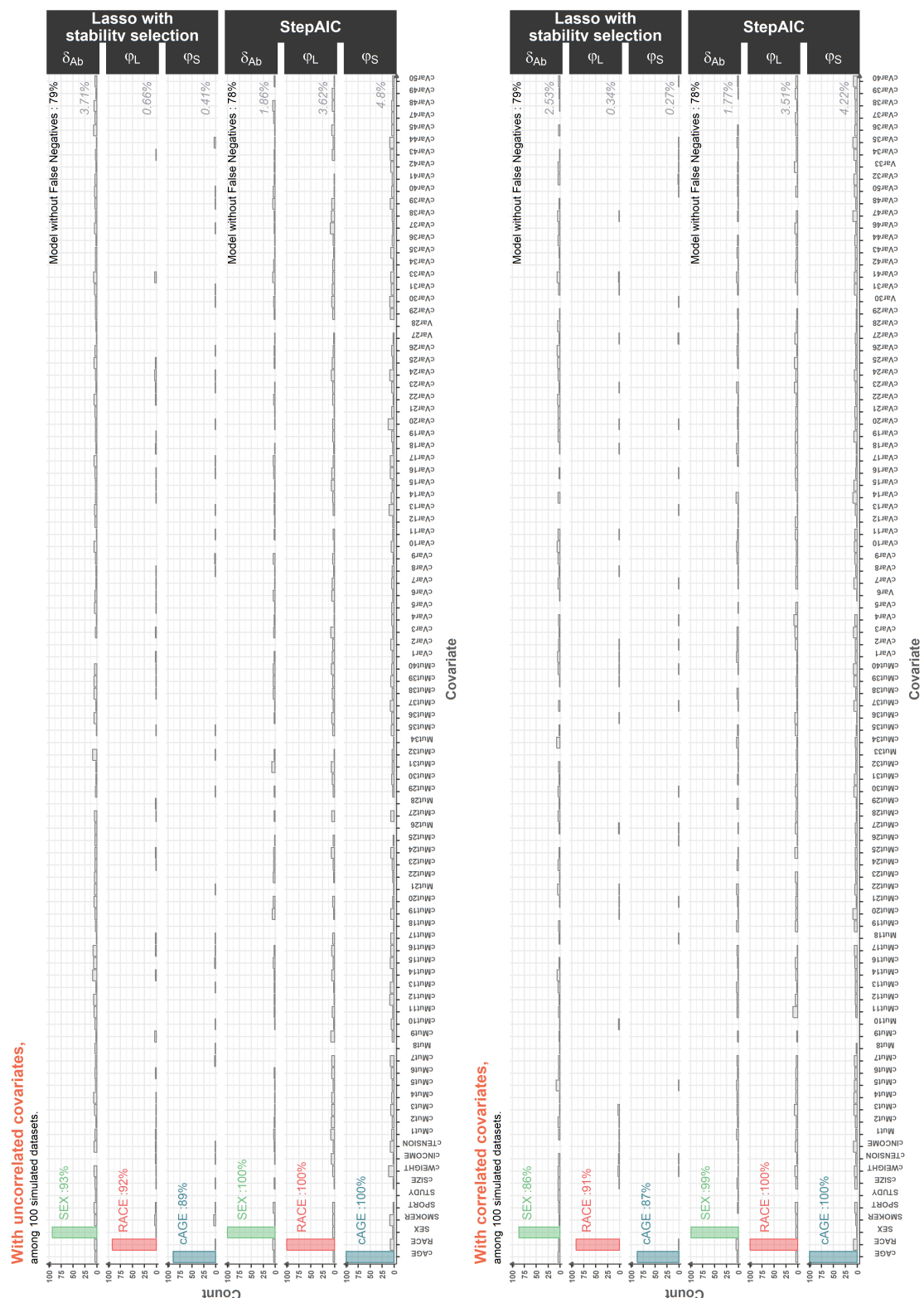
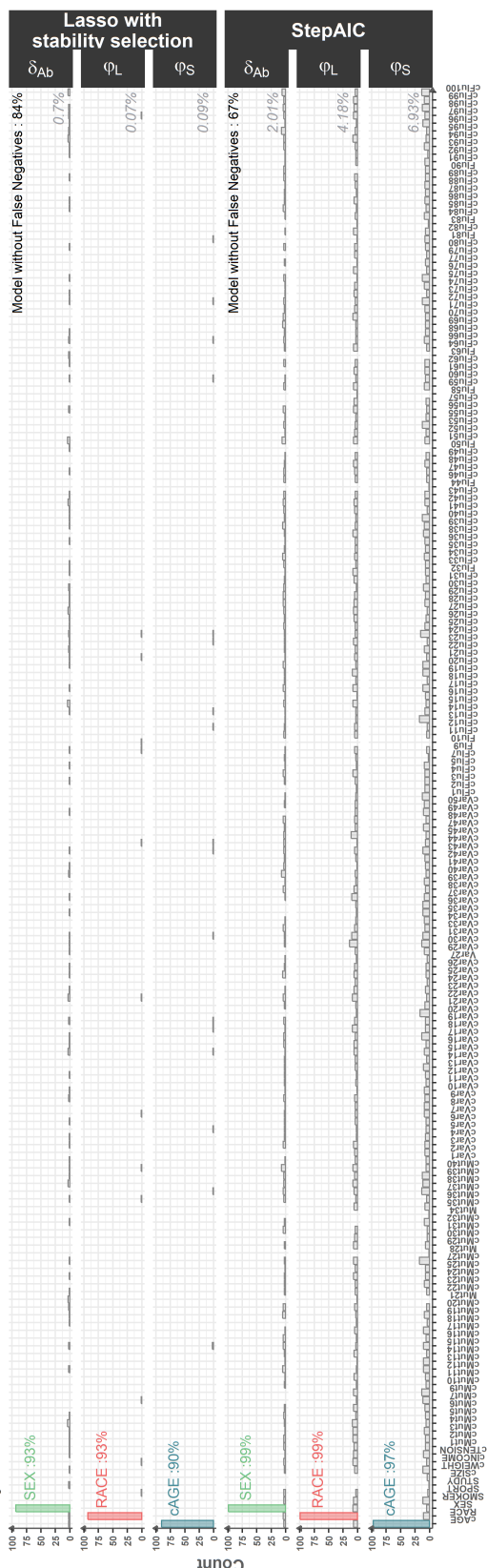


FIGURE 22 – Fréquence de sélection des différentes covariables avec chacun des paramètres, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations. Pour 100 covariables, 100 individus, avec le critère d'arrêt.

# Covariate Selection Count

With uncorrelated covariates,  
among 100 simulated datasets.



With correlated covariates,  
among 100 simulated datasets.

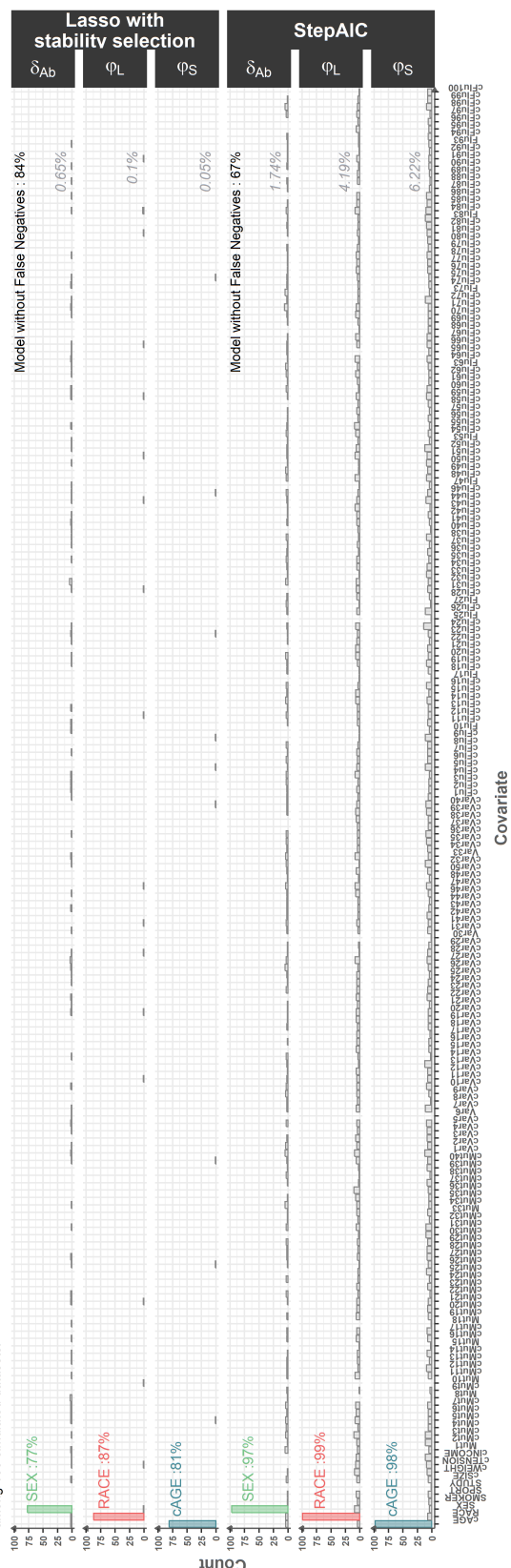


FIGURE 23 – Fréquence de sélection des différentes covariables avec chacun des paramètres, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations. Pour 200 covariables, 100 individus, avec le critère d'arrêt.

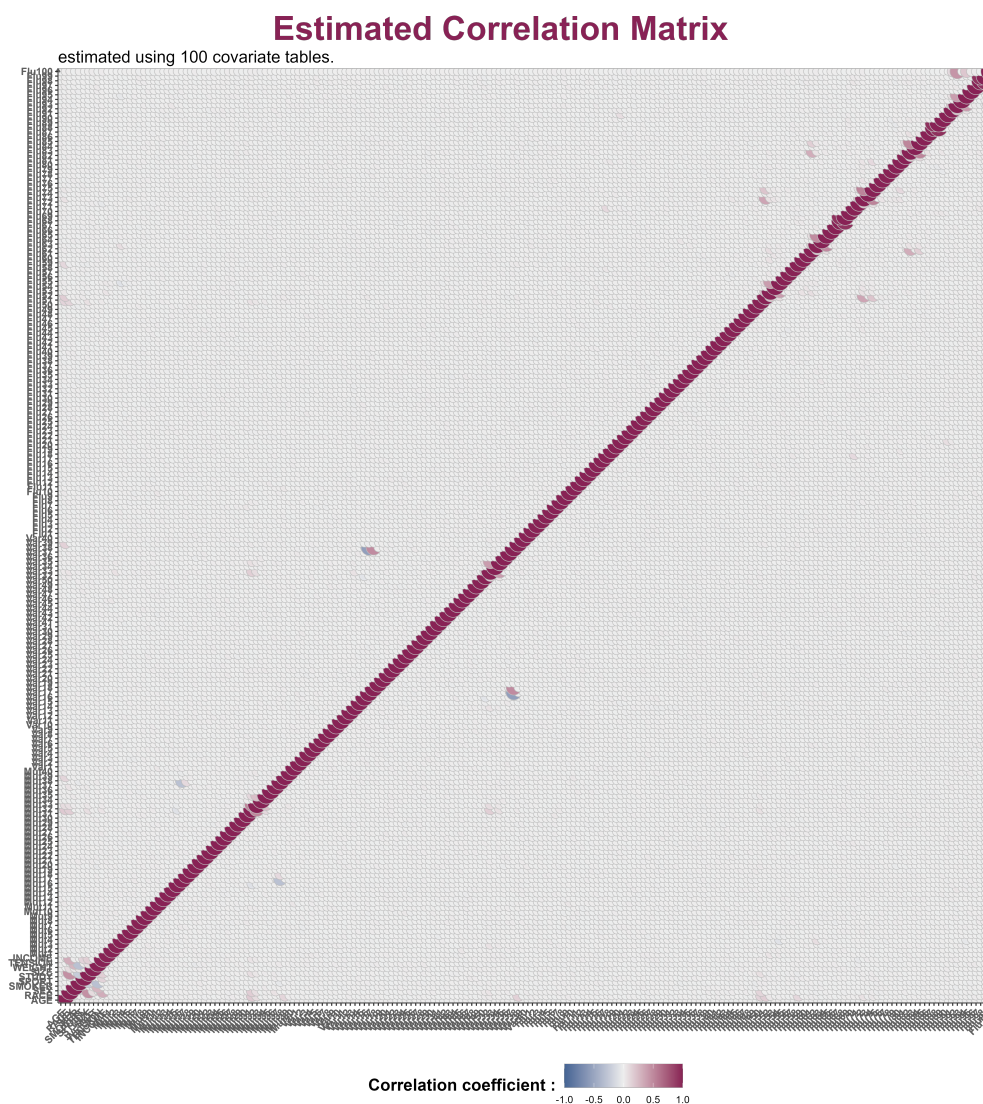


FIGURE 24 – Estimation des coefficients de corrélation de Spearman des 100 jeux de données créés, pour 200 covariables.

### Estimated Correlation Matrix

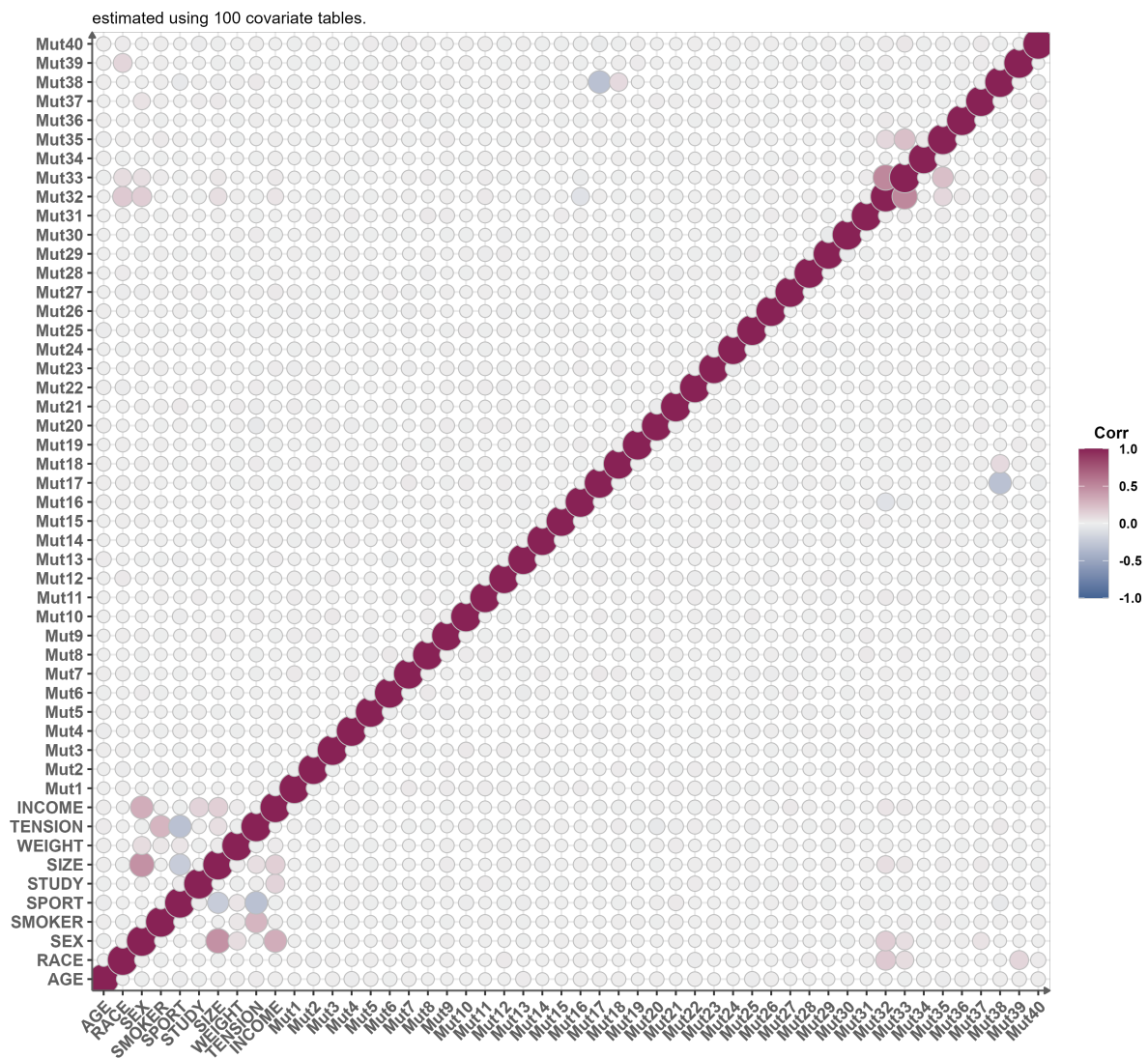
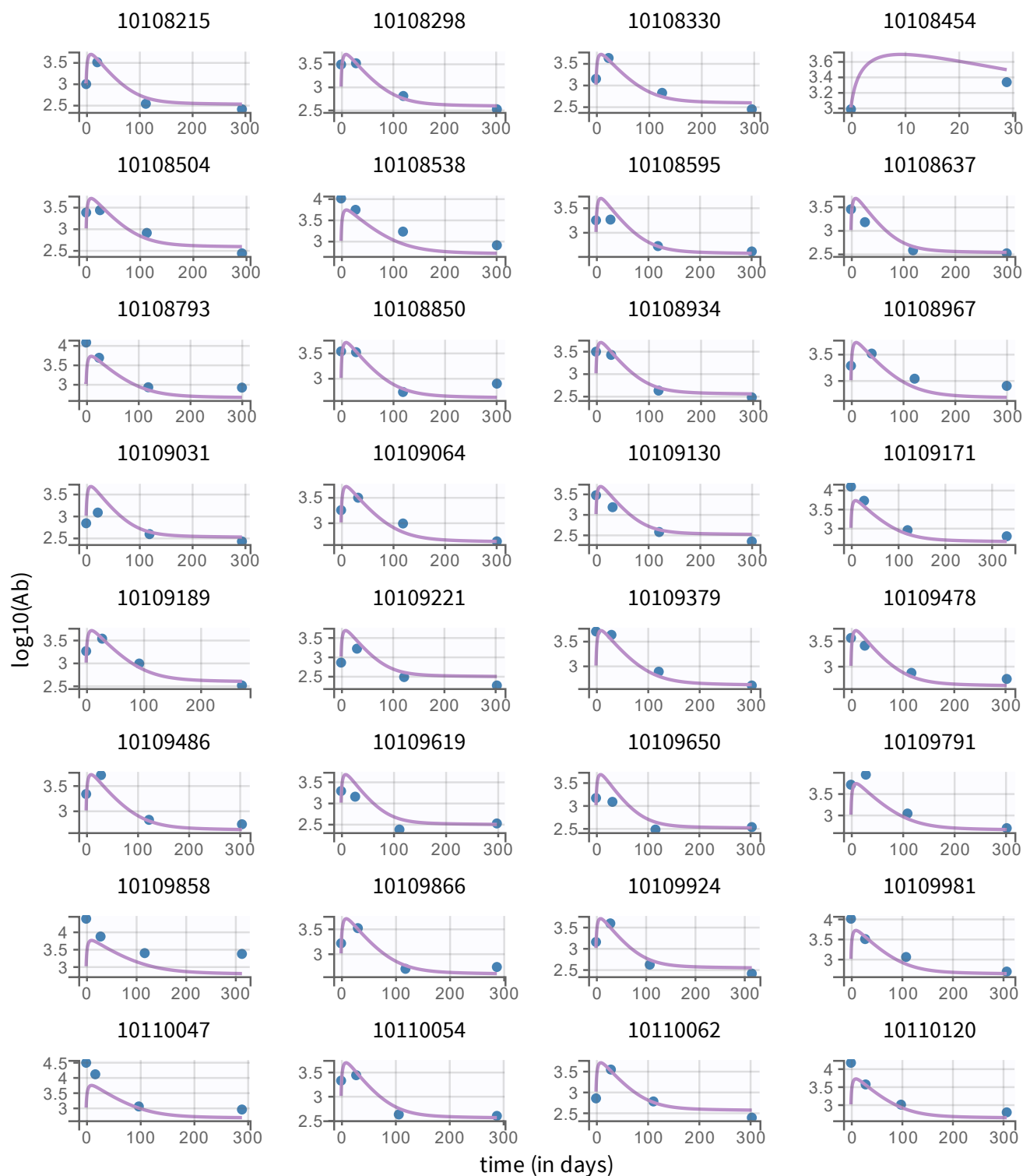
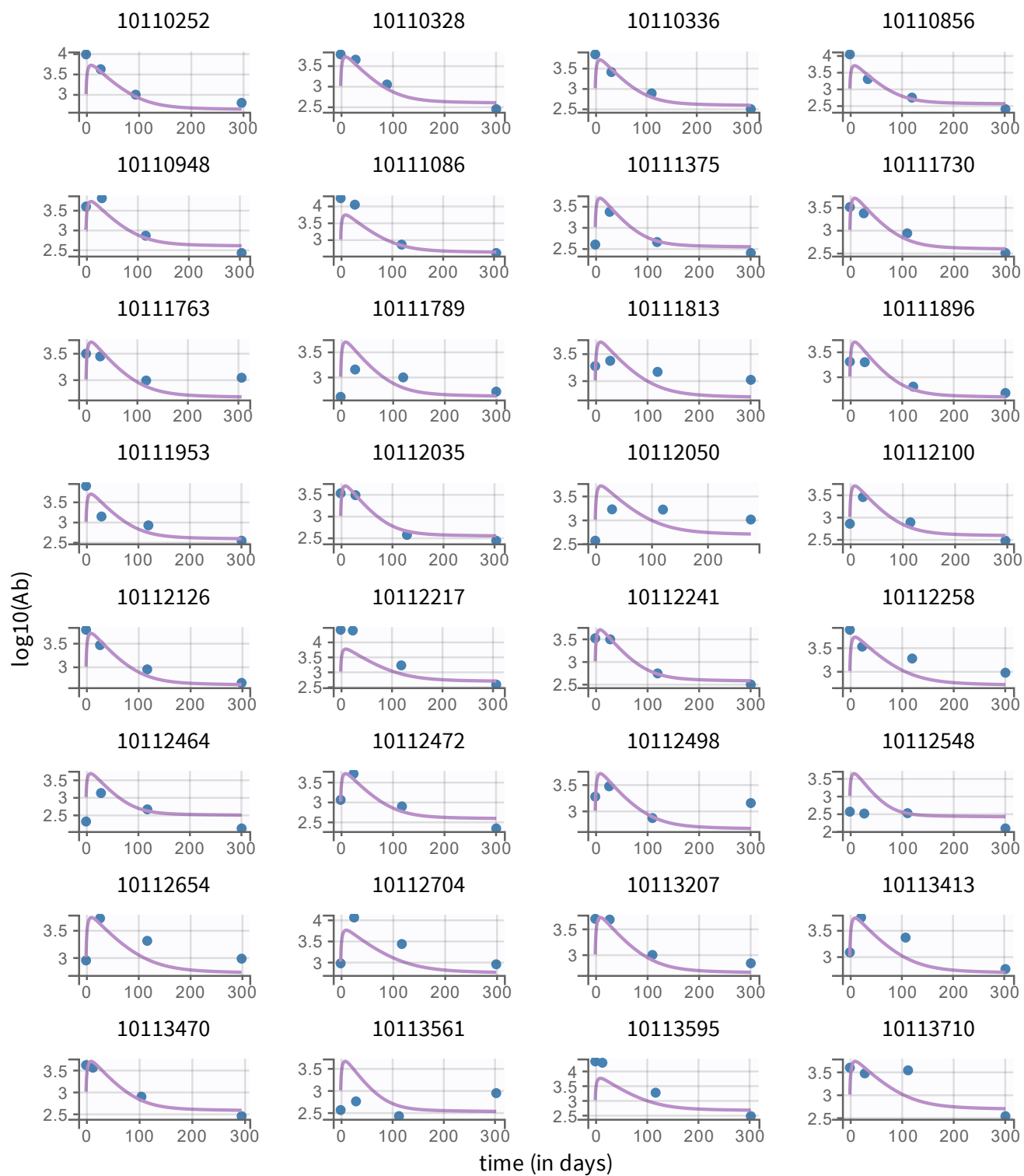


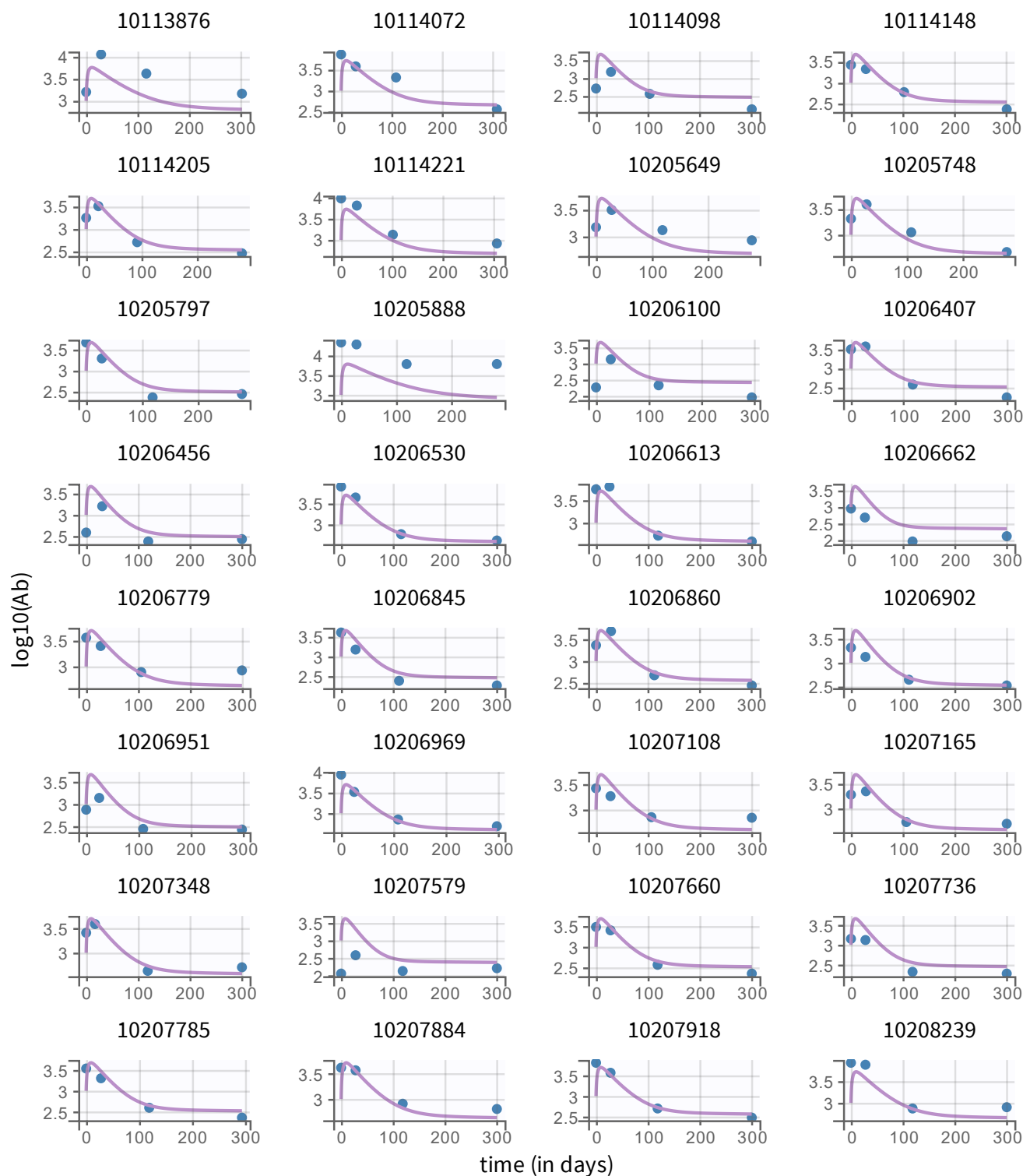
FIGURE 25 – Estimation des coefficients de corrélation de Spearman des 100 jeux de données créés, zommé sur les 50 premières covariables.

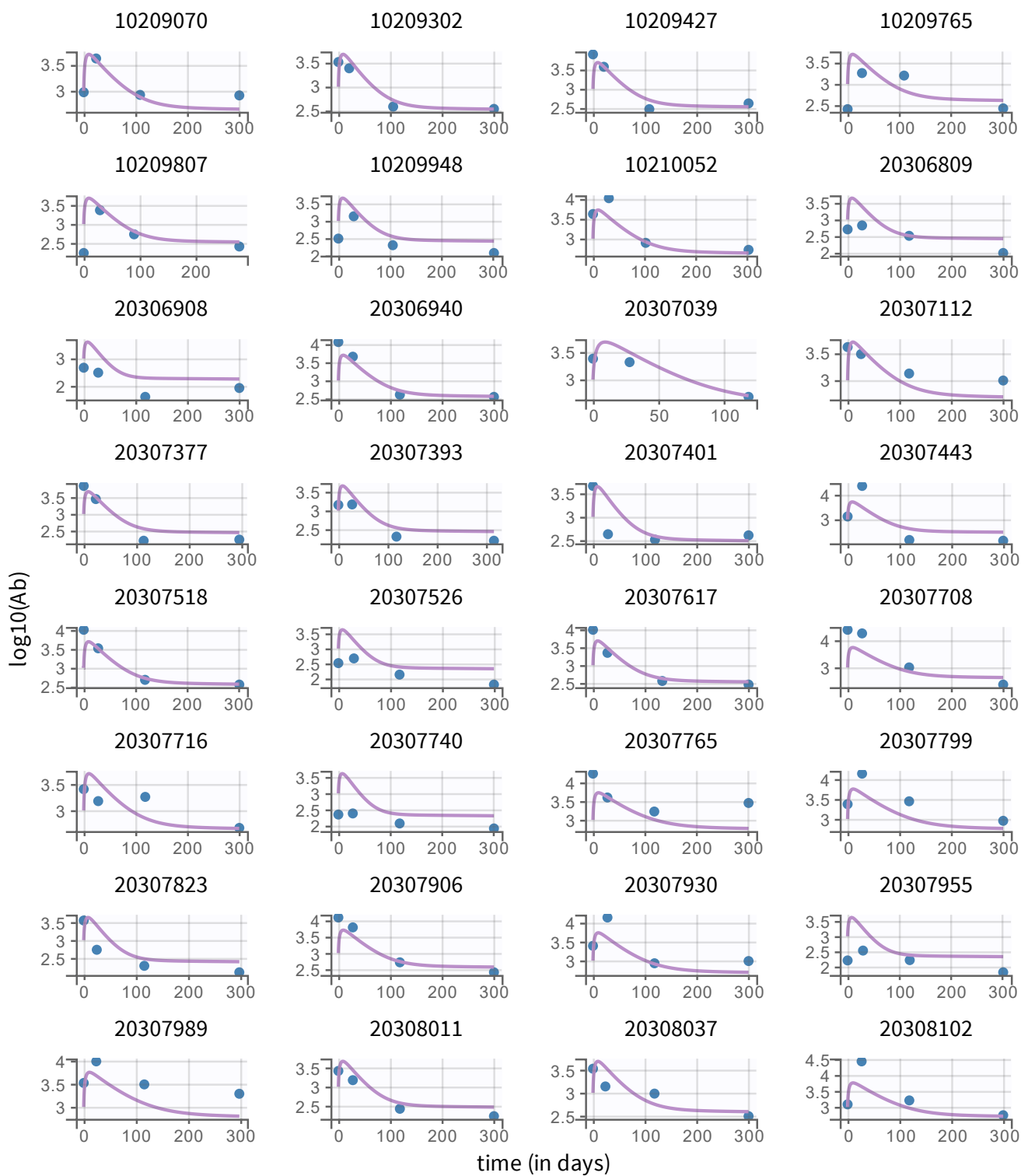
### C.6. Fits individuels



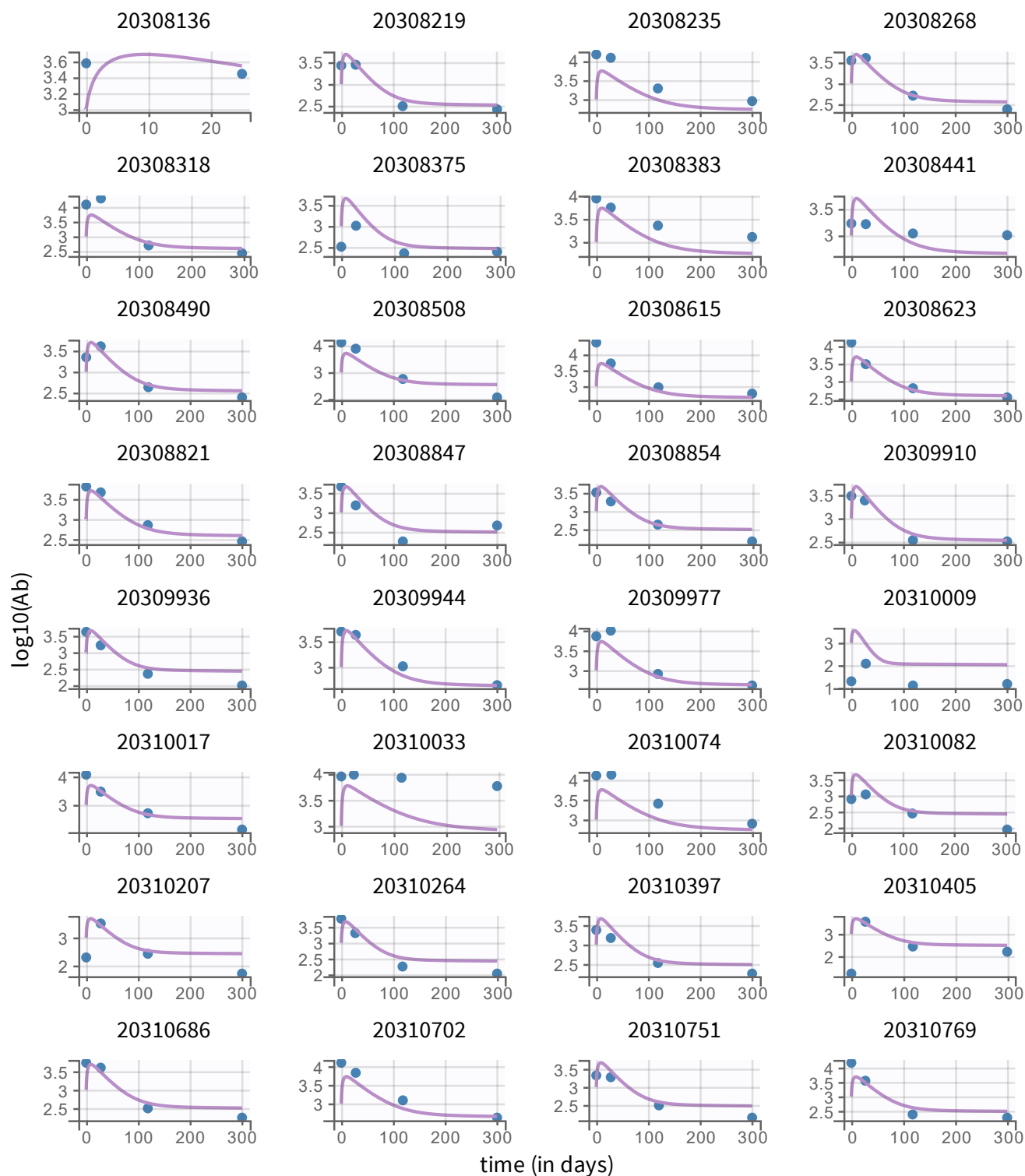


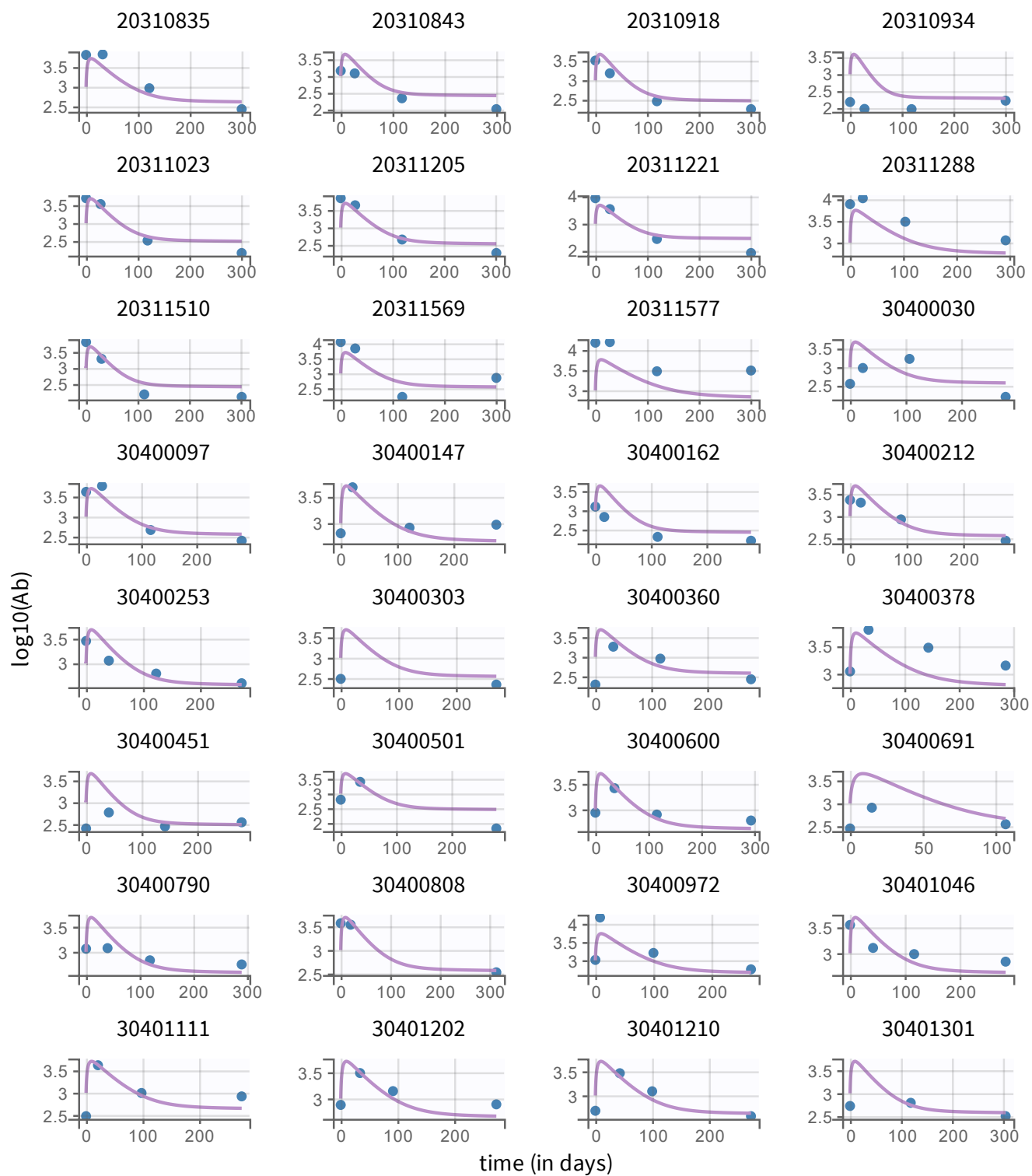


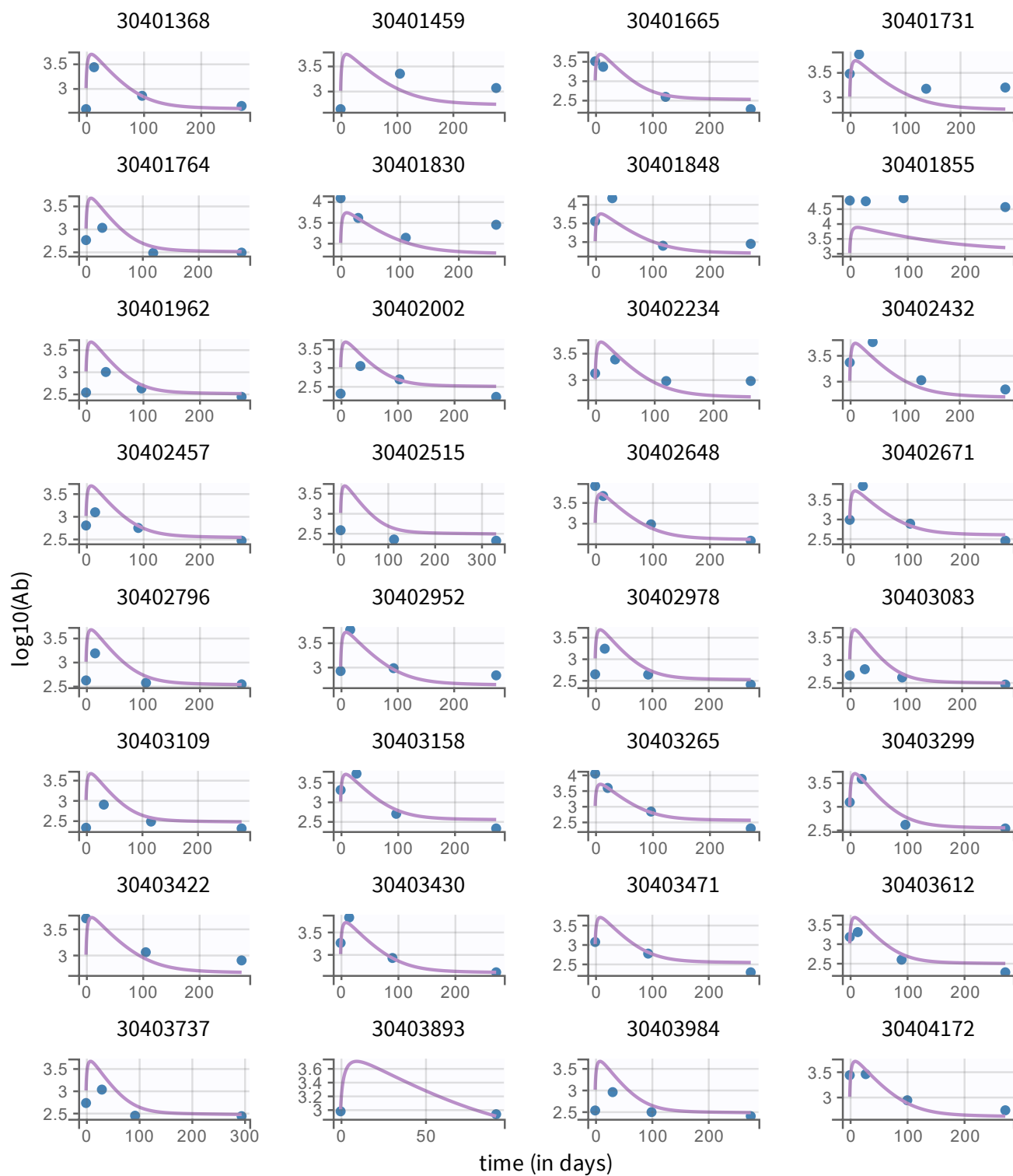


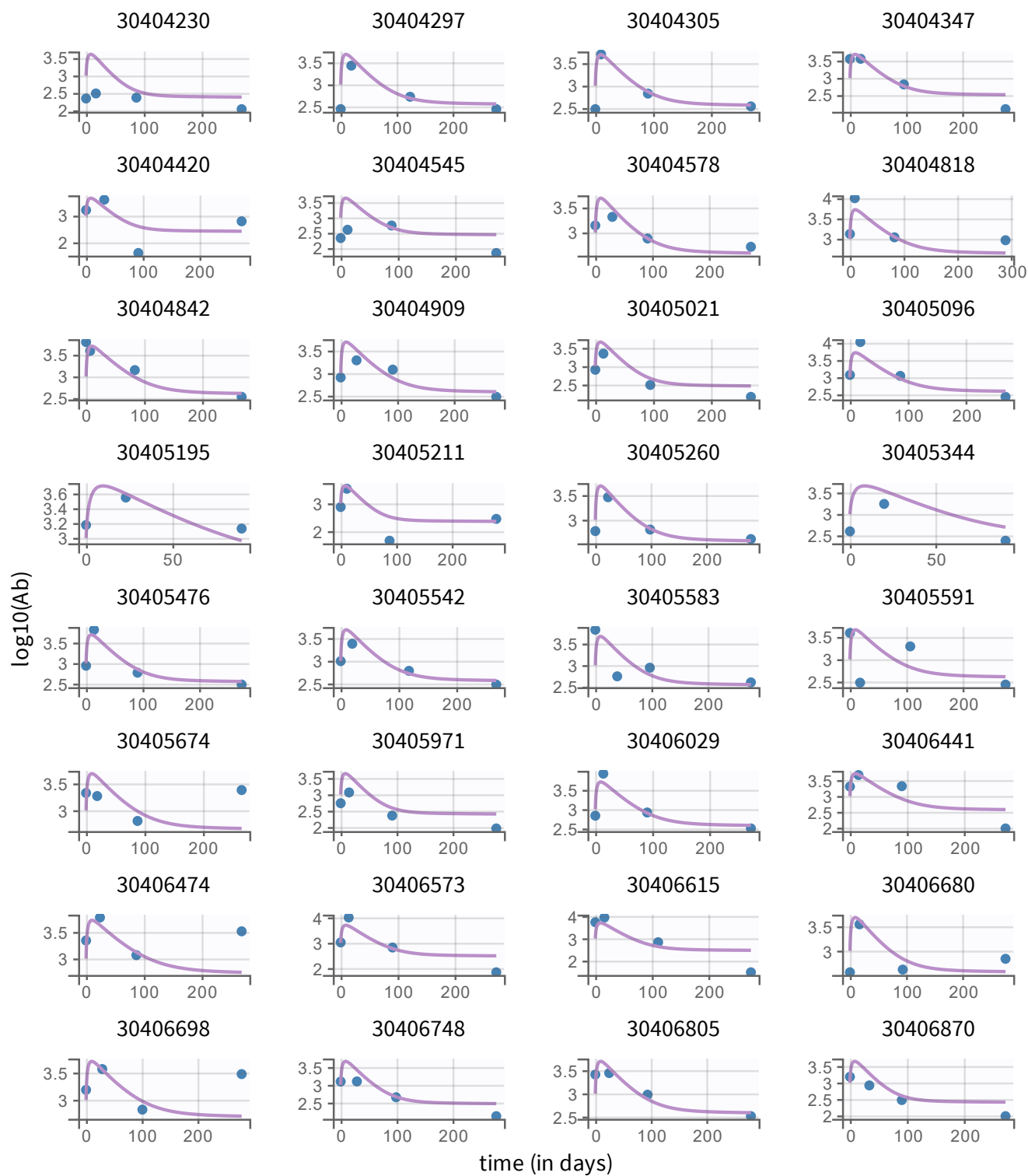


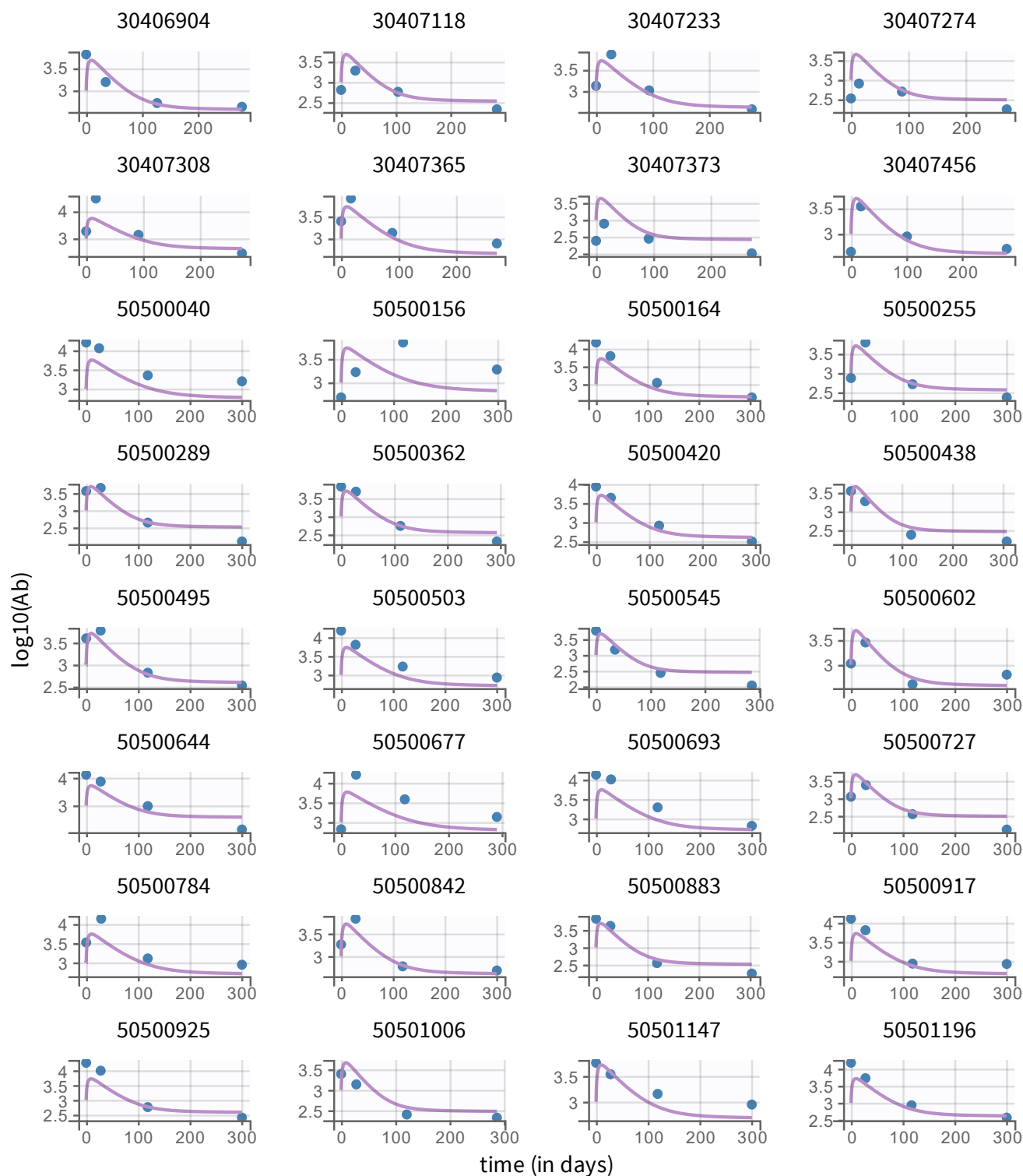












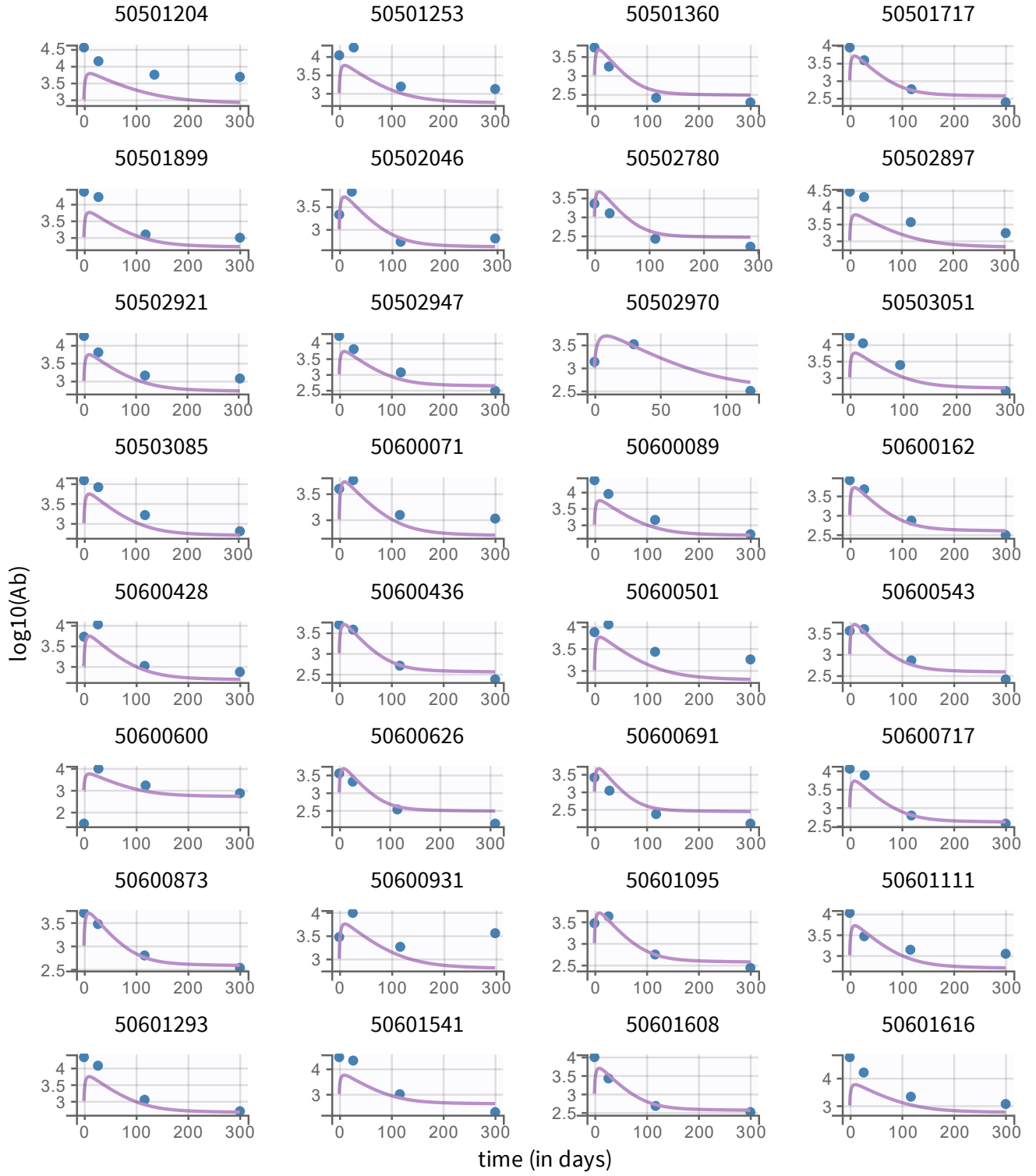
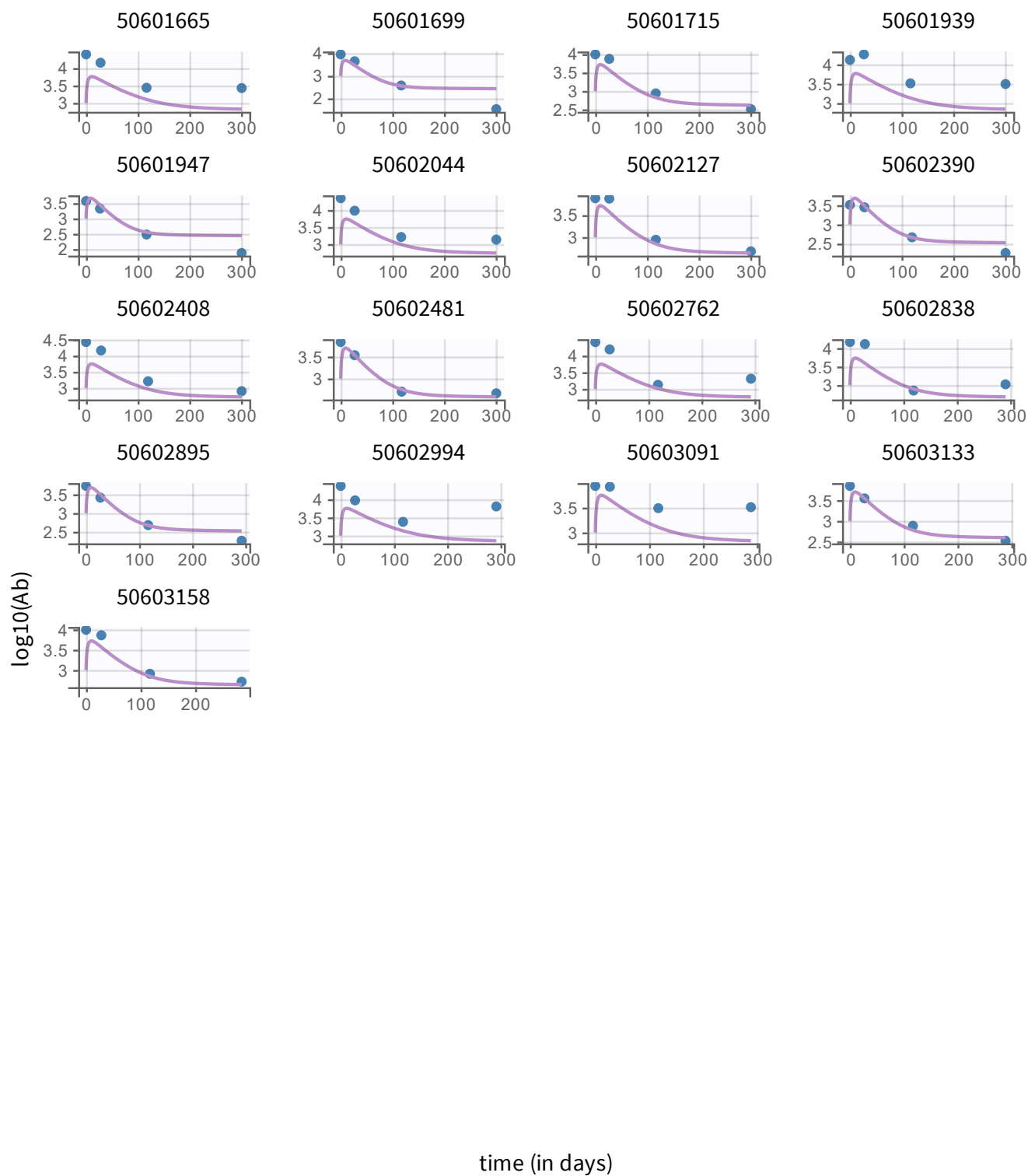


Figure 1: Individual fits



---

## D. Présentation et soumission du travail lié au stage et en vue de la thèse

---

### D.1. Poster pour la journée des jeunes chercheurs du BPH

A l'occasion de la seconde [journée des jeunes chercheurs du BPH](#), le 18 septembre 2023, j'ai présenté le poster présenté sur la page suivante. Il consiste en la méthode développée pendant mon stage et en préliminaire du sujet de thèse. A l'occasion de l'évènement, une récompense était attribué au meilleur poster, avec une bourse mobilité de 500€ pour effectuer une conférence, une école d'été ou autre évènement scientifique. J'ai eu la chance de recevoir cette récompense.

### D.2. Journée de Biostatistique - GDR de Toulouse

Les prochaines journées de Biostatistique organisées conjointement par le G.D.R. "Statistique et Santé", la Société française de biométrie et par le groupe "Biopharmacie et Santé" de la Société française de statistiques se tiendront les 16 et 17 novembre 2023 à l'Institut de Mathématiques de Toulouse, Université Paul Sabatier - Toulouse III. A cette occasion, je viens de soumettre l'abstract suivant, pour une présentation orale. Ces deux évènements me permettraient de me familiariser avec le milieu de la recherche et d'inscrire mon sujet de stage dans ma thèse.

#### **High-dimension Mechanistic Model Building using LASSO Approaches : Application to Ebola Vaccination**

##### **Contenu :**

Constructing non-linear mixed-effects models (NLMEM) deepens our comprehension of biological processes. Specifically, NLMEM facilitates the incorporation of inter-individual variability by parameterizing models at the individual level within a population framework. To do so, parameters combine fixed effects, capturing population-level relationships with covariates, and random effects, accounting for individual deviations. Estimation in NLMEM is achievable through maximum likelihood methods, such as the Stochastic Approximation Expectation-Maximization (SAEM-Dempster, 1977 ; Kuhn & Lavielle, 2005) algorithm. However, this approach is computationally intensive, and selecting covariates that define individual-level parameters cannot be done by comparing all possible models.

For the optimized construction of NLMEM, traditional methodologies rely on modified stepwise approaches (SCM-Jonsson, 1998 ; COSSAC-Ayral, 2021). Alternatively, the Stochastic Approximation for Model Building Algorithm (SAMBA-Prague & Lavielle, 2022) builds the covariate model on the



posterior realization of the parameters. Within a low-dimensional context, SAMBA efficiently and more rapidly constructs models by minimizing an information criterion. However, we aim to extend it for high-dimensional settings—such as those involving transcriptomic data. Initially, SAMBA employs a stepwise AIC algorithm for covariate selection. Our proposal integrates a multivariate LASSO approach, offering a more nuanced treatment of parameter correlations. This methodology incorporates a whitening step (Perrot-Dockès, 2018) and a stability selection process (Meinshausen & Bühlmann, 2010).

We validated our approach through simulations imitating the dynamics of the humoral immune response to an Ebola vaccine (Pasin, 2019). These simulations were replicated 100 times, each involving 100 individuals and 200 covariates. Remarkably, the False Discovery Rate for the proposed method was reduced by a factor of 10, while maintaining a similar False Negative Rate. This indicates enhanced control over the False Positive Rate. We applied our method using data from the Prevac/Prevac-UP trial (Prevac-UP Team, 2022), which compares two licensed vaccines for Ebola in Africa.

**Auteurs principaux :** GABAUT, Auriane (Université de Bordeaux, Inria, Inserm, Bordeaux Population Health Research Center, SISTM Team ; Vaccine Research Institute, Créteil, France) ; Mme PRAGUE, Mélanie (Université de Bordeaux, Inria, Inserm, Bordeaux Population Health Research Center, SISTM Team ; Vaccine Research Institute, Créteil, France)

**Orateur :** GABAUT, Auriane (Université de Bordeaux, Inria, Inserm, Bordeaux Population Health Research Center, SISTM Team ; Vaccine Research Institute, Créteil, France)

Déposé par Mlle GABAUT, Auriane le **lundi 2 octobre 2023**.

# High-dimension Mechanistic Model Building using LASSO Approaches :

## Application to Ebola Vaccination

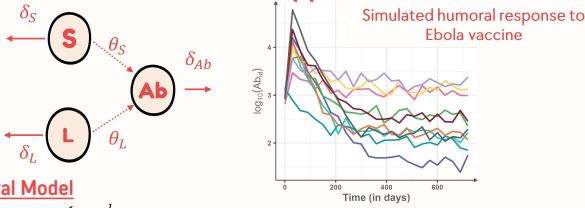
Auriane Gabaut, Mélanie Prague  
Inria Bordeaux Sud-Ouest, Inserm, Bordeaux Population Health Research Center, SISTM Team, UMR 1219, University of Bordeaux, Bordeaux, France  
This study was carried out in the framework of the University of Bordeaux's France 2030 program / RRI PHDS  
Contact : [auriane.gabaut@u-bordeaux.fr](mailto:auriane.gabaut@u-bordeaux.fr)

### Introduction

- Non-Linear Mixed Effects Model [1] allows to model complex biological processes such as response to vaccination.
- Understanding the determinant of vaccine response is key for vaccination strategy, it includes effect of sociodemographic covariates (low dimension) and transcriptomic data (high dimension).
- Data are collected during **Prevac-Up** trial [2], the immuno substudy gathers genetic expression of 95 participants.



### Non-Linear Mixed effects model



#### 1. Structural Model

$$\forall i \leq N, j \leq n_i, \begin{cases} \frac{d}{dt} Ab_i(t_{ij}) = \varphi_{S_i} e^{-\delta_S t_{ij}} + \varphi_{L_i} e^{-\delta_L t_{ij}} - \delta_{Ab_i} Ab_i(t_{ij}) \\ Ab_i(t_{i0}) = 0 \end{cases} = Ab_{0i}$$

$$\psi_i = (\varphi_{S_i}, \varphi_{L_i}, \delta_{Ab_i}, \delta_S, \delta_L), \quad h(\psi_i) = h(\psi_{pop}) + X_i \beta + \eta_i$$

#### 2. Statistical Model

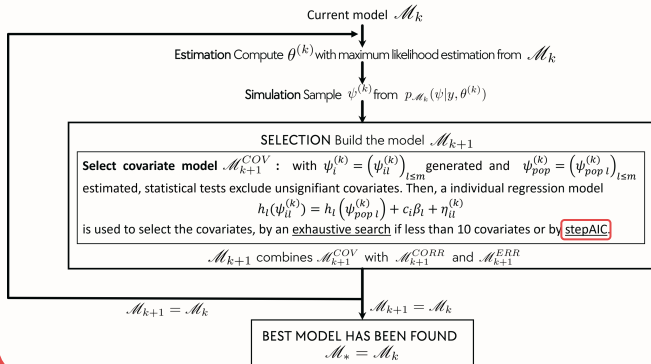
$$\forall i \leq N, \begin{cases} \eta_i^S \sim \mathcal{N}(0, \omega_S^2) \\ \eta_i^L \sim \mathcal{N}(0, \omega_L^2) \\ \eta_i^{Ab} \sim \mathcal{N}(0, \omega_{Ab}^2) \end{cases}$$

#### 3. Observation Model

$$\forall i \leq N, j \leq n_i, \begin{cases} Y_{ij} = \log_{10}(Ab_i(t_{ij})) + \varepsilon_{ij} \\ \varepsilon_i \sim \mathcal{N}(0, \Sigma = \sigma_{Ab}^2 I_{n_i}) \end{cases}$$

### SAMBAs [3]

Goal : to optimize the information criterion of a model.



### Method

#### Multivariate LASSO Selection within SAMBA

##### Multivariate LASSO [4,5]

With the model  $\mathcal{M}_k$  built at iteration  $k$ , and the estimated parameters  $\theta^{(k)}$ , we define :

$$h_l(\psi_{il}^{(k)}) = h_l(\psi_{pop}^{(k)}) + X_i \beta_l + \eta_{il}^{(k)}$$

It rewrites in a multivariate form:

$$\Psi^{(k)} = \left( h_l(\psi_{il}^{(k)}) \right)_{i \leq N, l \leq m} = \mathbf{1}_N \times \left( h_l(\psi_{pop}^{(k)}) \right)_{l \leq m} + X \beta + \eta^{(k)}$$

And, by whitening [5] the previous equation and vectorizing, we finally get the linear regression :

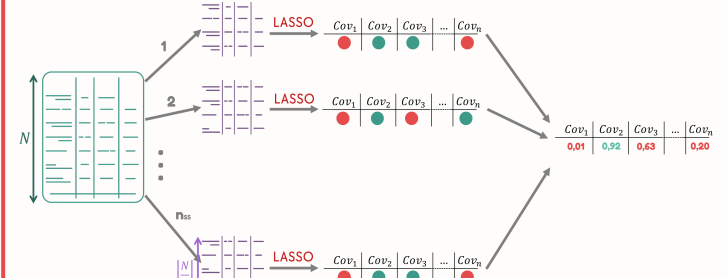
$$\mathbf{y}^{(k)} = \mathbf{y}_{pop}^{(k)} + X \mathbf{B} + \mathbf{E}$$

With :

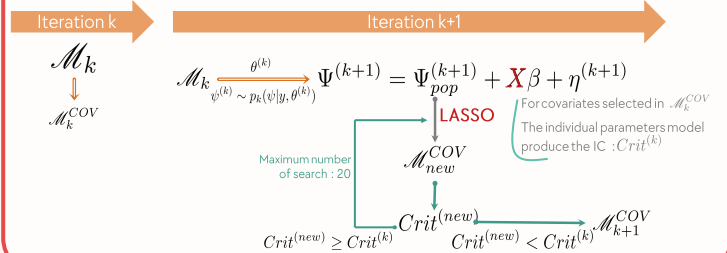
- $\mathbf{y}^{(k)} = \text{vec}(\Psi^{(k)} \times \Omega^{(k)^{-\frac{1}{2}}})$  and  $\mathbf{y}_{pop}^{(k)} = \text{vec}(\Psi_{pop}^{(k)} \times \Omega^{(k)^{-\frac{1}{2}}})$ , where  $\text{vec}$  is the operator vectorizing a matrix by its columns.
- $\mathbf{X} = (\Omega^{(k)^{-\frac{1}{2}}})^T \otimes X$ ;
- $\mathbf{B} = \text{vec}(\beta)$ ;
- $\mathbf{E} = \text{vec}(\eta^{(k)} \Omega^{(k)^{-\frac{1}{2}}})$ .

With the linear model obtain, we obtain a lasso estimator that is used to select covariates. The whitening step allows to consider the estimated correlation matrix  $\Omega^{(k)}$ .

##### Stability Selection [6]

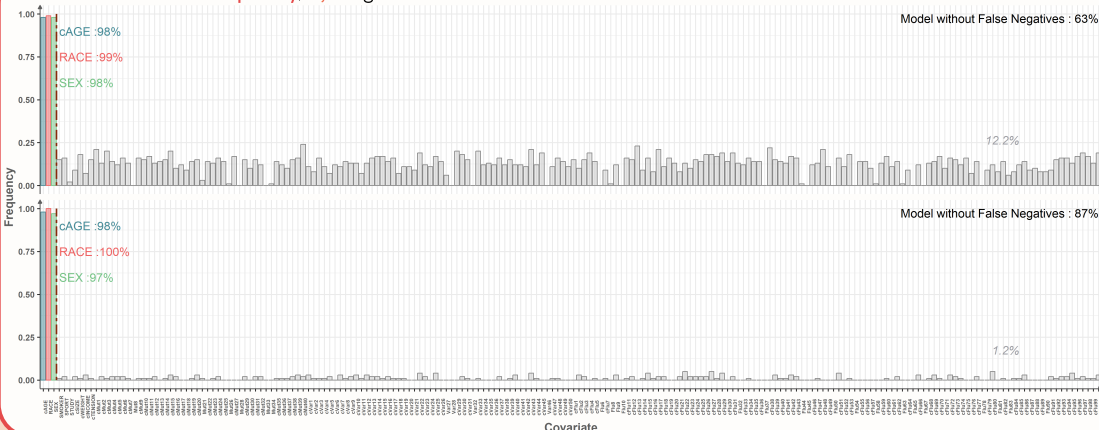


The final covariates selected are the one with a higher selection frequency (typically greater than a thresholds  $t_{SS} = 0,90$ ).

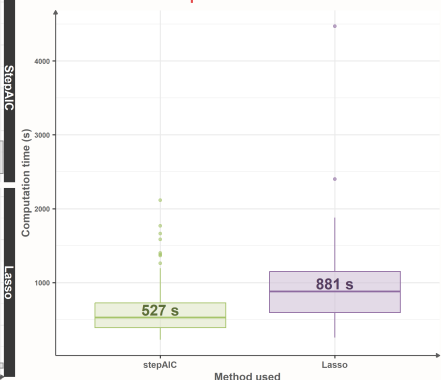


### Results

Covariate Selection Frequency, Among 100 simulated datasets with 100 individuals with 200 covariates.



#### Computation Time



### References

- [1] N. M. Laird et J. H. Ware, Biometrics, vol. 38, no 4, déc. 1982.
- [2] PREVAC Study Team, N Engl J Med, vol. 387, déc. 2022.
- [3] M. Prague et M. Lavielle, CPT Pharmacom & Syst Pharma, vol. 11, n° 2, févr. 2022.
- [4] R. Tibshirani, Journal of the Royal Statistical Society. Series B (Methodological), vol. 58, no 1, 1996.
- [5] M. Perrot-Dockès (Perrot-Dockès), « Méthodes régularisées pour l'analyse de données multivariées en grande dimension : théorie et applications. », PhD Thesis, 2019.
- [6] N. Meinshausen et P. Bühlmann, Journal of the Royal Statistical Society Series B: Statistical Methodology, vol. 72, no 4, sept. 2010.
- [7] C. Pasin et al., J Virol, vol. 93, no 18, sept. 2019
- [8] Lavielle, M. (2022). Rsmx: R Speaks 'Monolix'. R package version 5.0.1.
- [9] Friedman J, Tibshirani R, Hastie T (2010). Journal of Statistical Software, 33(1)

---

## Liste des références

---

ALEXANDRE, M., PRAGUE, M., MCLEAN, C., BOCKSTAL, V., DOUGUIH, M. et THIÉBAUT, R. (2023). Evaluatoin and prediction of the long-term humoral immune response induved by the two-dose heterologous Ad26.ZEBOC,MVA-BN-Filo vaccine regimen against ebola. draft manuscript, pre-submitted.

AYRAL, G., SI ABDALLAH, J.-F., MAGNARD, C. et CHAUVIN, J. (2021). A novel method based on unbiased correlations tests for covariate selection in nonlinear mixed effects models : The cossac approach. *CPT : Pharmacometrics & Systems Pharmacology*, 10(4) :318–329.

EFRON, B. et TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap*. Numéro 57 de Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, Boca Raton, Florida, USA.

GOLDFELD, K. et WUJCIAK-JENS, J. (2020). simstudy : Illuminating research methods through data generation. *Journal of Open Source Software*, 5(54) :2763.

SPEARMAN, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1) :72–101.

SVENSSON, R. J. et JONSSON, E. N. (2022). Efficient and relevant stepwise covariate model building for pharmacometrics. *CPT : Pharmacometrics & Systems Pharmacology*, 11(9) :1210–1222.

# Table des figures

,map size=0.9,widthh=

|      |                                                                                                                                                                     |     |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.1  | Lars Hörmander en 1969 . . . . .                                                                                                                                    | 58  |
| 4.2  | Lars Hörmander recevant la médaille Fields en 1962 . . . . .                                                                                                        | 59  |
| 6.1  | Les composants d'un système de R.L. [2]. . . . .                                                                                                                    | 106 |
| 6.2  | Image du tableau n°5 du jeu vidéo <i>Bubble Ghost</i> sorti sur Atari ST en 1987. Auteur du jeu : Christophe Andréani. . . . .                                      | 107 |
| 6.3  | Représentation schématique d'un neurone. source : <a href="https://culturesciencesphysique.ens-lyon.fr/">https://culturesciencesphysique.ens-lyon.fr/</a> . . . . . | 112 |
| 6.4  | Schéma d'un réseau de neurone. source : <a href="https://culturesciencesphysique.ens-lyon.fr/">https://culturesciencesphysique.ens-lyon.fr/</a> . . . . .           | 113 |
| 6.5  | Score total obtenu au fur et à mesure du temps par l'agent. . . . .                                                                                                 | 117 |
| 6.6  | Score moyen obtenu par l'agent au cours des $N$ derniers épisodes, avec $N = 50$ et $200$ . . . . .                                                                 | 118 |
| 6.7  | Score moyen obtenu par l'agent au cours des $N$ derniers épisodes, avec $N = 50$ et $200$ . . . . .                                                                 | 119 |
| 6.8  | Score total obtenu au fur et à mesure du temps par l'agent. . . . .                                                                                                 | 120 |
| 6.9  | Score total obtenu au fur et à mesure du temps par l'agent DQN et DDQN . . . . .                                                                                    | 121 |
| 6.10 | Score moyen obtenu par l'agent DQN et DDQN au cours des $N$ derniers épisodes, avec $N = 50$ et $200$ . . . . .                                                     | 121 |

|      |                                                                                                                                                                                            |     |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 7.1  | Schéma du modèle SIR considéré. . . . .                                                                                                                                                    | 133 |
| 7.2  | Graph du noyau d'Epanechnikov, pour $\delta = 1, 2$ et $10$ . . . . .                                                                                                                      | 141 |
| 7.3  | Système limite et simulation pour les paramètres $\lambda_1 = 0.6, \lambda_2 = 0.2, \lambda_3 = 0.2, c = 0.02$ . . . . .                                                                   | 142 |
| 7.4  | Données observées. . . . .                                                                                                                                                                 | 143 |
| 7.5  | Histogramme des différentes valeur prior simulées, pondérées par $W_i$ , avec la moyenne pondérée des tirages (en pointillet), et les paramètres des données observées (en tiret). . . . . | 143 |
| 7.6  | Simulation du processus pour $\hat{\lambda}_{1moy}, \hat{\lambda}_{2moy}, \hat{\lambda}_{3moy}, \hat{c}_{moy}$ . . . . .                                                                   | 144 |
| 7.7  | Simulation du processus pour $\hat{\lambda}_{1max}, \hat{\lambda}_{2max}, \hat{\lambda}_{3max}, \hat{c}_{max}$ . . . . .                                                                   | 144 |
| 8.1  | Protocole d'une analyse protéomique par spectrométrie de masse . . . . .                                                                                                                   | 149 |
| 8.2  | Densité des lois de Student pour différentes valeurs de $k$ . . . . .                                                                                                                      | 152 |
| 8.3  | Exemple de structure en arbre. . . . .                                                                                                                                                     | 159 |
| 8.4  | Expérience effectuée pour produire le premier jeu de données étudié. . . . .                                                                                                               | 163 |
| 8.5  | Répartition des données dans chaque condition, avant et après imputation. . . . .                                                                                                          | 164 |
| 8.6  | Répartition des données des filtrations gardées. . . . .                                                                                                                                   | 165 |
| 8.7  | Fonction de répartition empirique des p-valeurs selon $H_0$ ou $H_1$ avec la méthode sans seuil. . . . .                                                                                   | 166 |
| 8.8  | Fonction de répartition empirique des p-valeurs selon $H_0$ ou $H_1$ avec la méthode avec seuil. . . . .                                                                                   | 166 |
| 8.9  | Calculs des différentes bornes post hoc pour le premier jeu de données. . . . .                                                                                                            | 167 |
| 8.10 | Expérience effectuée pour produire le second jeu de données étudié. . . . .                                                                                                                | 168 |
| 8.11 | Répartition des valeurs mesurées selon plusieurs méthodes de prétraitement. . . . .                                                                                                        | 168 |
| 8.12 | Fonction de répartition empirique des p-valeurs selon la condition . . . . .                                                                                                               | 169 |
| 8.13 | Calcul des différentes bornes post hoc sur le second jeu de données. . . . .                                                                                                               | 170 |
| 8.14 | Calcul des différentes bornes post hoc sur le second jeu de données pour les plus petites p-valeurs. . . . .                                                                               | 170 |
| 8.15 | Calcul des différentes bornes, et de bornes par la méthode de $\lambda$ -calibration, sur le second jeu de données. . . . .                                                                | 171 |
| 8.16 | Calcul de $V_{\gamma}^{hyb}$ . . . . .                                                                                                                                                     | 172 |
| 8.17 | Calcul de $V_{\gamma}^{hyb,2}$ . . . . .                                                                                                                                                   | 173 |
| 8.18 | Comparaison des deux bornes $V_{\gamma}^{hyb,2}$ et $V_{\gamma}^{hyb}$ . . . . .                                                                                                           | 173 |

|      |                                                                                                                                                                                                                                                                                                                             |     |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 8.19 | Application de $V_{DKWM}^*$ dans le cas d'une sélection dans les données, les méthode "oracle" correspondent au vrai nombre de faux positifs obtenus sur l'ensemble $S_k$ obtenu par la méthode de rejet mentionné . . . . .                                                                                                | 174 |
| 9.1  | Trajectoires individuelles simulées de la quantité d'anticorps observée dans la réponse immunitaire humorale à un premier boost de vaccin contre Ebola. . . . .                                                                                                                                                             | 182 |
| 9.2  | Modèle de la réponse immunitaire humorale : les cellules sécrétrices d'anticorps Short-live (S) et Long-live (L) sécrètent respectivement aux taux $\theta_S$ et $\theta_L$ des anticorps (Ab). Ces entités biologiques se désagrègent respectivement aux taux $\delta_S$ , $\delta_L$ et $\delta_{Ab}$ . . . . .           | 185 |
| 9.3  | Trajectoire moyenne simulée de la dynamique des anticorps dans la réponse immunitaire humorale à un premier boost de vaccin contre Ebola. . . . .                                                                                                                                                                           | 186 |
| 9.4  | Trajectoires individuelles simulées de la dynamique des anticorps dans la réponse immunitaire humorale à un premier boost de vaccin contre Ebola. . . . .                                                                                                                                                                   | 186 |
| 9.5  | Vue d'ensemble de l'algorithme Stochastic Approximation for Model Building Algorithm (SAMBA), extrait de l'article [31] . . . . .                                                                                                                                                                                           | 190 |
| 9.6  | Estimation de la fonction de coût (en rouge) et la pénalité (en bleu) dans le cas du LASSO (à gauche) et de la régression ridge, avec une pénalisation $L^2$ (à droite). . . . .                                                                                                                                            | 193 |
| 9.7  | Trajectoires individuelles de la réponse immunitaire humorale 7 jours après un premier boost des participants à l'essai clinique PREVAC-Up[3] ; avec la trajectoire moyenne et les intervalles de confiance à 95% en chaque jour de visite théorique d'observation. La stratégie vaccinale est rHAd26/MVA-BN-Filo . . . . . | 197 |
| 9.8  | Design de l'essai clinique PREVAC-Up, avec l'exemple de la stratégie vaccinale avec comme prime Ad26.ZEBOV et comme boost MVA-BN-Filo. . . . .                                                                                                                                                                              | 197 |
| 9.9  | Fréquence de sélection des différentes covariables, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations, pour 100 covariables, 100 individus. . . . .                                                                                                                                  | 200 |
| 9.10 | Fréquence de sélection des différentes covariables, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations, pour 200 covariables, 100 individus. . . . .                                                                                                                                  | 201 |
| 9.11 | Fréquence de sélection des différentes covariables, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations, pour 100 covariables, 100 individus ; le test d'arrêt sur la décroissance du critère, ici le (BICc), est appliqué. . . . .                                                    | 202 |
| 9.12 | Fréquence de sélection des différentes covariables, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations pour 200 covariables, 100 individus ; le test d'arrêt sur la décroissance du critère, ici le (BICc), est appliqué. . . . .                                                     | 203 |
| 9.13 | Répartition du False Discovery Rate selon la méthode et la présence ou non de corrélations. Pour 100 covariables, 100 individus, à gauche et 200 covariables, 100 individus, à droite. . . . .                                                                                                                              | 204 |
| 9.14 | Temps d'exécution de l'algorithme de construction du modèle final selon la méthode de sélection utilisée, pour 200 covariables et 100 individus. . . . .                                                                                                                                                                    | 205 |

|      |                                                                                                                                                                                                                                                                                                                                                            |     |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 9.15 | Trajectoires individuelles de la réponse immunitaire humorale 7 jours après un premier boost des participants à la sous-étude immunologique de l'essai clinique PREVAC-Up[3] ; avec la trajectoire moyenne et les intervalles de confiance à 95% en chaque jour de visite théorique d'observation. La stratégie vaccinale est rHAd26/MVA-BN-Filo . . . . . | 206 |
| 10.1 | Représentation des dimensions courantes pour les données humorales, démographiques et de transcriptomiques. . . . .                                                                                                                                                                                                                                        | 212 |
| 10.2 | Simulations de la réponse immunitaire humorale au vaccin Ebola [9, 1]. On observe des trajectoires bruitées individuelles, et surtout, non linéaire ; les modèles mécanistes offrent la possibilité d'étudier efficacement ces données. . . . .                                                                                                            | 213 |
| 10.3 | Schéma de la production d'anticorps dans le cadre d'une réponse immunitaire humorale [9]                                                                                                                                                                                                                                                                   | 216 |
| 10.4 | Schéma du processus latent pour le modèle non-linéaire à effets mixtes de la réponse immunitaire. . . . .                                                                                                                                                                                                                                                  | 218 |
| 10.5 | Schéma réflectif du modèle d'observation . . . . .                                                                                                                                                                                                                                                                                                         | 219 |
| 10.6 | Schéma formatif du modèle d'observations . . . . .                                                                                                                                                                                                                                                                                                         | 220 |
| 10.7 | A gauche logo du projet Prevac-UP, essai vaccinale contre Ebola ; A droite, logo du Vaccine Research Institute (VRI), avec qui l'équipe SISTM travaille. . . . .                                                                                                                                                                                           | 221 |
| 10.8 | Gant du calendrier prévisionnel pour le projet de thèse. . . . .                                                                                                                                                                                                                                                                                           | 222 |
| 9    | Interrogation écrite n°22 . . . . .                                                                                                                                                                                                                                                                                                                        | 226 |
| 10   | Interrogation écrite n°23 . . . . .                                                                                                                                                                                                                                                                                                                        | 226 |
| 11   | Interrogation écrite n°24 . . . . .                                                                                                                                                                                                                                                                                                                        | 227 |
| 12   | Interrogation écrite n°25 . . . . .                                                                                                                                                                                                                                                                                                                        | 227 |
| 13   | Enoncé du DM . . . . .                                                                                                                                                                                                                                                                                                                                     | 228 |
| 14   | Programme de colle n°28 . . . . .                                                                                                                                                                                                                                                                                                                          | 229 |
| 15   | Programme de colle n°29 . . . . .                                                                                                                                                                                                                                                                                                                          | 230 |
| 16   | Programme de colle n°30 . . . . .                                                                                                                                                                                                                                                                                                                          | 231 |
| 17   | Vue d'ensemble des algorithmes Stepwise Covariate Modelling (SCM) et Stepwise Covariate Modelling amélioré (SCM+), extrait de l'article [6]. . . . .                                                                                                                                                                                                       | 237 |
| 18   | Schéma de la procédure de construction de modèle de covariables COnditional Sampling use for Stepwise Approach based on Correlation tests (COSSAC). La notation LL représente la log-vraisemblance multiplié par $-2$ . Figure extraite de [2] . . . . .                                                                                                   | 238 |
| 19   | Estimation de l'erreur quadratique moyenne pour différente valeur de $\lambda$ pour des données simulées du modèle (MOD AB). . . . .                                                                                                                                                                                                                       | 240 |

|    |                                                                                                                                                                                                                                                |     |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 20 | Fréquence de sélection des différentes covariables avec chacun des paramètres, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations. Pour 100 covariables, 100 individus. . . . .                          | 243 |
| 21 | Fréquence de sélection des différentes covariables avec chacun des paramètres, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations. Pour 200 covariables, 100 individus. . . . .                          | 244 |
| 22 | Fréquence de sélection des différentes covariables avec chacun des paramètres, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations. Pour 100 covariables, 100 individus, avec le critère d'arrêt. . . . . | 245 |
| 23 | Fréquence de sélection des différentes covariables avec chacun des paramètres, selon la méthode de sélection utilisé dans SAMBA et la présence, ou non, de corrélations. Pour 200 covariables, 100 individus, avec le critère d'arrêt. . . . . | 246 |
| 24 | Estimation des coefficients de corrélation de Spearman des 100 jeux de données créés, pour 200 covariables. . . . .                                                                                                                            | 247 |
| 25 | Estimation des coefficients de corrélation de Spearman des 100 jeux de données créés, zommé sur les 50 premières covariables. . . . .                                                                                                          | 248 |



# Listre des tableaux

,map size=0.9,widthh=

|     |                                                                                                                                                                                                                                                                                                                                 |     |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 6.1 | Hyperparamètres utilisés . . . . .                                                                                                                                                                                                                                                                                              | 115 |
| 8.1 | $V_{DKWM}^*$ calculé sur $(S_k)_{k \in \mathcal{K}}$ , correspondant au résultat des procédures de rejet appliquées aux $k$ plus petites p-valeurs obtenues, l'oracle correspond au vrai nombre de faux positifs $ S_k \cap \mathcal{H}_1 $ dans cet ensemble $S_k$ , et le False Discovery Proportion (FDP) de $S_k$ . . . . . | 175 |
| 9.1 | Résultat de l'algorithme de stabilité sélection, avec $n_{SS} = 1000$ , sur un jeu de données simulés. Les fréquences de sélection inférieur à 0.05 sont indiqués en clair ■, et celle supérieure à 0.95 sont indiqués en foncé ■. Les fréquences soulignés sont celles des bonnes associations variables-paramètres . . . . .  | 199 |
| 9.2 | Valeur des différents paramètres du modèle (MOD AB STR) estimés par le logiciel Monolix.                                                                                                                                                                                                                                        | 206 |
| 1   | Valeur des paramètres $\psi$ utilisée pour les simulations, d'après [1]. . . . .                                                                                                                                                                                                                                                | 241 |
| 2   | Procédé de génération des différentes loi par la fonction <code>randomCovariate</code> implémenté . .                                                                                                                                                                                                                           | 242 |