

Dynamique markovienne ternaire cyclique sur graphes et quelques applications en biologie mathématique

Mémoire

Vincent Painchaud

Sous la direction de:

Nicolas Doyon
Patrick Desrosiers

Résumé

La modélisation de phénomènes biologiques qui impliquent un très grand nombre d'unités pose toujours un défi. De nombreux modèles présentent une vision globale de la dynamique moyenne du phénomène sous la forme d'un système d'équations différentielles ordinaires. C'est le cas notamment du modèle de Wilson–Cowan, qui décrit l'activité qui se propage dans un réseau de neurones biologiques. Une limite importante de ce modèle est qu'il néglige d'éventuelles corrélations entre les états de différents neurones. L'objectif premier de ce mémoire est ainsi de le généraliser afin de décrire de telles corrélations. On veut aussi mieux en comprendre les fondements mathématiques et les liens qu'il a avec des modèles semblables utilisés en épidémiologie et en écologie.

Pour s'attaquer à ce problème, on construit une chaîne de Markov en temps continu qui décrit l'évolution des états des nœuds d'un graphe, et qui peut ainsi modéliser un phénomène biologique d'un point de vue microscopique. Étant donné le très grand nombre de nœuds que comporte le graphe, ce modèle microscopique est difficile à analyser. À partir du processus stochastique, on obtient alors par un moyennage un système d'équations différentielles ordinaires afin de décrire la dynamique sur le graphe d'un point de vue macroscopique. Deux applications de cette méthode sont alors présentées : l'une en épidémiologie et l'autre en neurosciences. On se concentre particulièrement sur l'application en neurosciences, qui permet de décrire la dynamique d'un réseau de neurones biologiques et de généraliser le modèle de Wilson–Cowan. En effet, on arrive à proposer deux nouveaux systèmes qui sont des extensions de ce modèle, puisqu'elles permettent de considérer des corrélations entre les états de différents neurones. On présente finalement un exemple dans lequel le comportement dynamique de l'une de ces extensions est plus près du comportement du processus stochastique que celui du modèle de Wilson–Cowan.

Abstract

Modeling biological phenomena that involve a very large number of individual units is always a challenge. In this context, many models consist in a system of ordinary differential equations that gives an overview of the mean dynamics of a phenomenon. Among these is the Wilson–Cowan model, which describes the activity of a biological neural network. An important weakness of this model is that it neglects all possible correlations between the states of different neurons. The main goal of this thesis is to generalize Wilson–Cowan’s model to describe such correlations. We also seek to get a better understanding of its mathematical foundations, as well as its links with other models used in epidemiology and ecology.

To tackle this problem, we construct a continuous-time Markov chain to describe the evolution of the states of the nodes of a large graph. Such a process can then model a biological phenomenon from a microscopic point of view. Since the size of the graph is very large, this microscopic model is hard to analyze. Hence, from the stochastic process, we use an averaging method to obtain a system of ordinary differential equations which describes the dynamics on the graph from a macroscopic point of view. We show two applications of this method : one in epidemiology and the other in neuroscience. We focus on the application in neuroscience, which leads to a description of the dynamics a biological neural network and generalizes Wilson–Cowan’s model. Indeed, we introduce two new systems which are extensions of this model since they can describe correlations between the states of different neurons. Finally, we present an example where the behavior of the stochastic process is closer to the dynamical behavior of one of the extensions than that of Wilson–Cowan’s model.

Table des matières

Résumé	ii
Abstract	iii
Table des matières	iv
Liste des figures	vi
Table des notations	viii
Remerciements	xii
Introduction	1
1 Systèmes dynamiques et probabilités	5
1.1 Notions de théorie des systèmes dynamiques	5
1.1.1 Équations différentielles ordinaires	5
1.1.2 Solutions fermées	8
1.1.3 Bifurcations de Hopf	10
1.2 Notions de théorie des probabilités	13
1.2.1 Espaces de probabilité et variables aléatoires	13
1.2.2 Indépendance et espérance conditionnelle	16
1.2.3 Espaces produit et noyaux de transition	18
1.2.4 Processus stochastiques	21
1.2.5 Chaînes de Markov à temps continu homogène	23
1.3 Problèmes de fermeture de moments	25
2 Dynamique markovienne ternaire cyclique sur graphes aléatoires	28
2.1 Graphes aléatoires	29
2.1.1 Graphes aléatoires pondérés	29
2.1.2 Paramètres et populations	30
2.2 Dynamique markovienne ternaire cyclique	31
2.2.1 Intuition sur la dynamique	31
2.2.2 Construction du processus de Markov	33
2.3 Réduction à un système d'équations différentielles ordinaires	38
2.4 Description macroscopique de la dynamique	43
2.5 Le problème de fermeture de moments	45

3 Application aux modèles épidémiologiques compartimentaux	48
3.1 Les modèles compartimentaux	48
3.1.1 Le modèle SIR classique	49
3.1.2 Le modèle SIRS	51
3.2 Modélisation d'une épidémie par une dynamique MTC	51
3.2.1 Modélisation de l'épidémie	52
3.2.2 Passage à la description macroscopique de la dynamique	53
3.2.3 Remarques sur les suppositions qui mènent au système (3.12)	55
3.3 Fermetures de moments au-delà de l'approximation de champ moyen	56
3.3.1 L'approximation des paires	57
3.3.2 L'approximation de champ moyen de deuxième ordre	60
3.3.3 Bref survol d'autres méthodes	62
3.4 Applicabilité à d'autres modèles compartimentaux	62
4 Application aux réseaux de neurones biologiques	64
4.1 Le modèle de Wilson–Cowan	65
4.1.1 Bref historique	65
4.1.2 Dérivation du modèle	66
4.1.3 Généralisations	68
4.1.4 Liens avec d'autres modèles	68
4.2 Modélisation de l'activité d'un réseau de neurones biologiques par une dynamique MTC	68
4.2.1 Modélisation du réseau de neurones biologiques	69
4.2.2 Passage à la description macroscopique de la dynamique	71
4.2.3 Réduction au modèle de Wilson–Cowan	73
4.3 Effet de l'état réfractaire sur le modèle de Wilson–Cowan	74
4.3.1 Quelques résultats généraux	75
4.3.2 Étude d'un exemple à une population	77
4.3.3 Étude d'un exemple à deux populations	81
4.3.4 Conclusion	84
4.4 Approximation de champ moyen de deuxième ordre	85
4.4.1 La fermeture de moments	86
4.4.2 Étude du cas à une population	89
4.4.3 Exemple à deux populations	104
4.4.4 Conclusion	106
4.5 Nouvelle solution au problème de fermeture de moments	106
4.5.1 Approximation de l'espérance d'une fonction de forme sigmoïdale	107
4.5.2 Fermeture de moments	118
4.5.3 Quelques résultats numériques	122
4.5.4 Conclusion	126
Conclusion	128
Bibliographie	131
A Dérivées des seconds moments	140
B PopNet	147

Liste des figures

1.1	Portrait de phase du système de Lotka–Volterra (1.5).	9
1.2	Portraits de phase du système (1.7) pour $\alpha = -1$	11
1.3	Portraits de phase du système (1.7) pour $\alpha = 1$	12
1.4	Diagrammes de bifurcation pour les bifurcations de Hopf sous-critique et supercritique du système (1.7).	12
2.1	Transitions permises entre les trois états.	32
3.1	Transitions entre les trois compartiments dans le modèle SIR.	49
3.2	Trois solutions du modèle SIR (3.3).	50
3.3	Transitions entre les trois compartiments dans le modèle SIRS.	51
4.1	Transitions entre les trois états d'un neurone.	70
4.2	Exemple de portrait de phase du système de champ moyen (4.17) pour une seule population.	76
4.3	Solutions du système de champ moyen (4.17) et du système réduit (4.19) pour les paramètres (4.31) et la condition initiale (4.32).	78
4.4	Portrait de phase du système de champ moyen (4.17) pour les paramètres (4.31).	78
4.5	Valeurs propres de la matrice jacobienne du système combiné (4.34) selon ε	80
4.6	Deux angles de vue du diagramme de bifurcation du système combiné (4.34) selon ε	80
4.7	Comparaison entre les solutions des systèmes (4.17) et (4.19) et une trajectoire du processus de Markov construit à la section 4.2 pour les paramètres (4.36) et l'état initial (4.37).	82
4.8	Valeurs propres de la matrice jacobienne du système combiné (4.39) selon ε	83
4.9	Diagramme de bifurcation du système combiné (4.39) selon ε	84
4.10	Solutions du système (4.55) pour les paramètres (4.121) et l'état initial (4.122), pour trois valeurs d'activités entrante externe.	102
4.11	Trajectoires possibles du processus de Markov construit à la section 4.2 pour les paramètres (4.121) et l'état initial (4.122), pour trois valeurs d'activité entrante externe.	103
4.12	Solutions du système (4.54) pour les paramètres (4.123) et l'état initial (4.124).	105
4.13	Trajectoire possible du processus de Markov construit à la section 4.2 pour les paramètres (4.123) et l'état initial (4.124).	106
4.14	Fonction de répartition d'une loi logistique.	109
4.15	Exemples de changements de facteur d'échelle dans des fonctions de répartition de lois logistiques.	111

4.16 Solutions du système (4.181) pour l'état initial (4.122) et pour les paramètres (4.121), pour trois valeurs d'activité entrante externe.	123
4.17 Solution du système (4.181) pour les paramètres (4.123) et l'état initial (4.124).	123
4.18 Pour les paramètres (4.183) et l'état initial (4.184), solutions du système de champ moyen (4.17) et du système (4.181), et statistiques obtenues à partir de trajectoires du processus stochastique sous-jacent.	124
4.19 Portrait de phase du système de champ moyen (4.17) pour les paramètres (4.183).	125
4.20 Projections des portraits de phase du système (4.181) dans le plan \mathcal{A} - \mathcal{R} pour des variances de 0, $1/1000$ et $2/1000$	126

Table des notations

Notations générales

$\mathbb{1}_E$	Fonction indicatrice de l'ensemble E : $\mathbb{1}_E(x) = 1$ si $x \in E$ et $\mathbb{1}_E(x) = 0$ si $x \notin E$.
2^E	Ensemble puissance de E .
E^T	Ensemble des fonctions de T vers E .
$ E $	Cardinal de E .
\cdot	Indique une dérivée, notamment dans le contexte d'une équation différentielle.
$\langle \cdot, \cdot \rangle$	Produit scalaire usuel.
$\mathcal{B}(E)$	σ -algèbre de Borel sur E , qui est supposé être un espace topologique.
$\mathcal{C}^k(E)$	Ensemble des fonctions k fois continûment dérivables sur E .
Df	Dérivée d'une fonction $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$; $Df(x)$ est la matrice jacobienne de f en x .
\mathbb{E}	Espérance par rapport à une mesure de probabilité \mathbb{P} .
$\mathbb{E}[Y x]$	$\mathbb{E}[Y \pi_E](x, y)$ pour $(x, y) \in E \times F$, où Y est \mathcal{F} -mesurable sur un espace de probabilité $(E \times F, \mathcal{E} \otimes \mathcal{F}, \mathbb{P})$ et où $\pi_E: E \times F \rightarrow E$ note la projection sur E . Voir proposition 1.13 et les remarques suivantes.
\mathbb{N}	Ensemble des nombres entiers naturels, excluant zéro.
\mathbb{N}_0	Ensemble des nombres entiers naturels, incluant zéro.
\mathbb{P}	Mesure de probabilité.

Notations relatives à une dynamique MTC générale

A_t^J	Fraction des nœuds de la population J dans l'état 1 au temps t . Voir (2.47).
$\mathbb{A}_J(t)$	Espérance de A_t^J . Voir (2.50).
a_J	Fonction $[0, 1]^{2n} \rightarrow [0, \infty)$ qui donne le taux de transition moyen $0 \mapsto 1$ des nœuds de la population J dans l'état 0. Voir (2.59).
a_j	Fonction $\Theta \times E \rightarrow [0, \infty)$ qui donne le taux de transition $0 \mapsto 1$ pour le nœud j .
b_J	Fonction $[0, 1]^{2n} \rightarrow [0, \infty)$ qui donne le taux de transition moyen $1 \mapsto i$ des nœuds de la population J dans l'état 1. Voir (2.59).
b_j	Fonction $\Theta \times E \rightarrow [0, \infty)$ qui donne le taux de transition $1 \mapsto i$ pour le nœud j .

c_J	Fonction $[0, 1]^{2n} \rightarrow [0, \infty)$ qui donne le taux de transition moyen $i \mapsto 0$ des nœuds de la population J dans l'état i . Voir (2.59).
c_j	Fonction $\Theta \times E \rightarrow [0, \infty)$ qui donne le taux de transition $i \mapsto 0$ pour le nœud j .
$C_{UV}^{JK}(t)$	Covariance entre U_t^J et V_t^K pour deux populations J et K du graphe, où U et V remplacent A, R, S , ou d'autres fonctions linéaires de ces variables.
(E, \mathcal{E})	Espace d'états ternaires sur un graphe. Voir définition 2.9.
$E_{UV}^{JK}(t)$	Espérance du produit $U_t^J V_t^K$ pour deux populations J et K du graphe, où U et V remplacent A, R, S , ou d'autres fonctions linéaires de ces variables.
J, K, \dots	Populations d'un graphe.
J_t^ξ	Ensemble des nœuds de la population J qui se trouve dans l'état $\xi \in \{0, 1, i\}$ au temps t . Voir (2.53).
N	Nombre de nœuds d'un graphe.
n	Nombre de populations d'un graphe.
(Ω, \mathcal{F})	Espace $(E, \mathcal{E})^{[0, \infty)}$ des trajectoires possibles du processus de Markov qui décrit l'état d'un graphe.
\mathcal{P}	Division d'un graphe en populations.
\mathbb{P}	Mesure de probabilité sur l'espace produit $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F})$. Voir corollaire 2 du théorème 2.3.
\mathbb{P}^ϑ	Mesure de probabilité sur (Ω, \mathcal{F}) pour un $\vartheta \in \Theta$ donné.
R_t^J	Fraction des nœuds de la population J dans l'état i au temps t . Voir (2.47).
$\mathcal{R}_J(t)$	Espérance de R_t^J . Voir (2.50).
S_t^J	Fraction des nœuds de la population J dans l'état 0 au temps t . Voir (2.47).
$\mathcal{S}_J(t)$	Espérance de S_t^J . Voir (2.50).
$(\Theta, \mathcal{G}, \mu)$	Espace des paramètres sur un graphe aléatoire muni d'une dynamique MTC.
W	Matrice aléatoire de poids.
W_{jk}	Variable aléatoire sur $(\Theta, \mathcal{G}, \mu)$ qui représente le poids du lien du nœud k vers le nœud j .
X_t	Variable aléatoire sur (Ω, \mathcal{F}) qui représente l'état d'un graphe au temps t .
X_t^j	Variable aléatoire sur (Ω, \mathcal{F}) qui représente l'état du nœud j au temps t .
Y_t	État macroscopique d'un graphe au temps t .
Y_t^J	État macroscopique (A_t^J, R_t^J) de la population J au temps t .

Notations relatives à une dynamique MTC appliquée en épidémiologie

0	État d'un individu susceptible.
1	État d'un individu infecté.
i	État d'un individu retiré.
β_{JK}	Taux $\beta_{JK} = K W_{JK}$ auquel un individu de la population J a des contacts avec les individus de la population K .
γ	Paramètre d'ordre 1 sur le graphe dont une composante γ_j représente le taux de guérison de l'individu j .
γ_J	Espérance de γ_j si $j \in J$.
I_t^J	Fraction des individus infectés de la population J au temps t ; autre notation pour A_t^J plus près de la notation usuelle du modèle SIR.
$\mathcal{I}_J(t)$	Espérance de I_t^J ; autre notation pour $\mathcal{A}_J(t)$ plus près de la notation usuelle du modèle SIR.
W	Matrice aléatoire de poids dont une composante W_{jk} représente le taux auquel les individus j et k ont des interactions qui mènent à l'infection de j par k .
W_{JK}	Espérance de W_{jk} si $j \in J$ et $k \in K$.

Notations relatives à une dynamique MTC appliquée en neurosciences

0	État d'un neurone sensible.
1	État d'un neurone actif.
i	État d'un neurone réfractaire.
α	Paramètre d'ordre 1 sur le graphe dont une composante α_j représente le taux d'activation du neurone j s'il reçoit au moins l'excitation de seuil.
α_J	Espérance de α_j si $j \in J$.
β	Paramètre d'ordre 1 sur le graphe dont une composante β_j représente le taux de désactivation du neurone j .
β_J	Espérance de β_j si $j \in J$.
B_t^J	Activité totale entrante dans la population J au temps t . Voir (4.12).
$\mathcal{B}_J(t)$	Espérance de B_t^J , donnée par $\sum_{K \in \mathcal{P}} c_{JK} \mathcal{A}_K(t)$.
c_{JK}	Poids de connexion qui représente l'activité que la population K envoie à la population J si tous ses neurones sont actifs.
\mathcal{D}_n	Ensemble des états macroscopiques du réseau de neurones. Voir (4.23).

$\mathcal{D}_n^{\text{wc}}$	Sous-ensemble de \mathcal{D}_n où $\mathcal{R}_J = \frac{\beta_J}{\gamma_J} \mathcal{A}_J$ pour chaque J . Voir (4.26).
\mathcal{D}_n^+	Domaine physiologique. Voir (4.67).
F_{θ_J}	Fonction de répartition des seuils θ_j pour $j \in J$.
γ	Paramètre d'ordre 1 sur le graphe dont une composante γ_j représente le taux auquel le neurone j passe de réfractaire à sensible.
γ_J	Espérance de γ_j si $j \in J$.
Q_J	Activité externe au réseau entrante à chaque neurone de la population J .
Q_j	Activité externe au réseau entrante au neurone j .
θ	Paramètre d'ordre 1 sur le graphe dont une composante θ_j représente le seuil d'activation du neurone j .
W	Matrice aléatoire de poids dont une composante W_{jk} représente l'activité que le neurone k envoie au neurone j lorsqu'il s'active.

Remerciements

Ce mémoire de maîtrise est l'aboutissement de plus de deux ans de travail durant lesquels j'ai été entouré par une foule de gens.

Je tiens tout d'abord à remercier mes directeurs de recherche, Patrick et Nicolas, pour leur soutien considérable depuis mes toutes premières expériences en recherche, durant mon baccalauréat. Toutes les discussions intéressantes que nous avons partagées m'ont énormément aidé, autant dans mon travail de recherche que dans mon développement comme apprenti-chercheur. Je me compte chanceux d'avoir pu travailler avec des superviseurs aussi disponibles et encourageants qu'eux. Je dois également les remercier de m'aider à comprendre la valeur de mon travail, que j'ai tendance à ne pas savoir reconnaître. J'espère sincèrement que notre collaboration n'est pas terminée.

Il me faut aussi adresser un remerciement particulier à Patrick pour m'avoir fait découvrir en premier ce que sont en fait les mathématiques. Je crois bien que sans son cours de Physique mathématique I, je ne serais pas en train d'écrire ces lignes.

Je remercie également tout le groupe Dynamica de m'avoir accueilli malgré ma nouvelle allégeance aux mathématiques. Ce fut un plaisir de pouvoir partager un peu de mon projet de recherche avec d'autres étudiants. Je tiens à remercier en particulier Béatrice, qui m'accompagne de près depuis le baccalauréat, pour son amitié irremplaçable. Merci d'avoir partagé avec moi l'aventure de la maîtrise, malgré la distance forcée.

J'ai aussi reçu énormément de soutien à l'extérieur du monde académique, en particulier de ma famille. Je veux d'abord remercier mes frères et sœurs : Laure, Anne et Émile, pour toutes les niaiseries, les farces et les rires. Vous êtes une équipe formidable dont je suis heureux et chanceux de faire partie. Je remercie aussi du fond du cœur mes parents pour leur présence constante et leur support sans bornes, dont ils semblent toujours refuser de comprendre l'importance.

Finalement, je tiens à souligner l'appui inconditionnel de Geneviève, qui m'accompagne dans les hauts et les bas de tous les jours. Particulièrement en cette période pandémique, ma motivation et mon moral auraient été énormément amoindris sans sa présence, son sourire et son écoute.

Introduction

Plusieurs phénomènes biologiques sont le résultat d'interactions complexes entre une multitude d'unités, d'individus ou de cellules. De bons exemples sont les mouvements de populations animales étudiés en écologie, la propagation d'une épidémie dans une population, ou encore l'activité échangée par des neurones dans un réseau. Ces exemples sont très semblables du point de vue d'un mathématicien qui voudrait les modéliser : on s'intéresse dans chaque cas à un grand nombre d'unités qui sont toutes interreliées, et qui peuvent se trouver dans quelques états distincts. Pour étudier de tels phénomènes en biologie mathématique, deux approches distinctes sont souvent utilisées [76].

D'une part, il est possible de modéliser la structure par un graphe dont les nœuds représentent les unités étudiées — par exemple des individus ou des cellules — et dont les liens représentent les relations entre elles — par exemple des relations sociales ou des synapses. Pour modéliser le phénomène biologique, on peut supposer que pour chacune de ces unités, le passage d'un état à un autre est aléatoire. Il est alors possible de construire un processus stochastique qui modélise la dynamique voulue d'un point de vue microscopique, au sens où les interactions individuelles entre chaque paire d'individus sont toutes décrites. On peut s'imaginer que ce genre de modèle est assez près de la réalité, mais très difficile à étudier analytiquement.

D'autre part, il est possible d'étudier les fractions d'unités qui sont dans chacun des états possibles, en adoptant un point de vue global pour observer le phénomène. Par exemple, afin d'étudier une épidémie, on pourrait simplement s'intéresser aux proportions x et y d'individus qui sont malades et d'individus qui ne le sont pas. Pour modéliser la dynamique, on peut ensuite supposer que les unités passent d'un état à un autre à certains taux donnés. En reprenant l'exemple de l'épidémie, on pourrait supposer¹ que les individus malades guérissent à un taux a et que les individus qui ne sont pas malades le deviennent à un taux b . Il est alors possible de construire un système d'équations différentielles ordinaires : avec l'exemple de l'épidémie on trouve que $\dot{x} = -ax + by$ et que $\dot{y} = -by + ax$. Un tel système décrit le phénomène d'un point de vue macroscopique, au sens où seule une vue d'ensemble de la situation est obtenue, sans détails sur les interactions entre les individus eux-mêmes. Ce genre de modèle est alors plus simple à étudier, mais plus approximatif.

Ces deux approches, qui semblent très différentes a priori, peuvent en fait souvent être re-

1. Évidemment, on veut seulement illustrer l'idée générale d'une telle modélisation ; les suppositions données ne mènent pas à un très bon modèle épidémiologique !

liées mathématiquement. En effet, un modèle macroscopique présenté sous la forme d'équations différentielles est typiquement construit pour représenter le comportement moyen ou attendu d'une dynamique stochastique décrite par un modèle microscopique. Un tel modèle macroscopique est souvent appelé *modèle de champ moyen* [60]. De façon plus générale, le passage d'un modèle microscopique vers un modèle macroscopique peut se faire en définissant des variables aléatoires qui représentent un état macroscopique du système. À partir du processus stochastique qui décrit le phénomène du point de vue microscopique, il est alors possible de décrire l'évolution des moments de ces variables aléatoires par des équations différentielles. Ceci fait survenir un problème de fermeture de moments : en arrivant à exprimer l'évolution de certains moments de façon autonome, on obtient un système dynamique qui décrit le phénomène d'un point de vue macroscopique.

Des exemples typiques qui présentent une telle dichotomie entre des versions microscopiques et macroscopiques sont les modèles compartimentaux étudiés en épidémiologie, reliés aux célèbres travaux [57] de Kermack et McKendrick. En effet, ils ont à l'origine été construits comme des modèles macroscopiques, et sont encore souvent présentés ainsi, comme dans [7, 46, 70]. Or, ils ont aussi été reformulés comme des processus de Markov qui décrivent les états de chacun des individus de la population, notamment dans [1, 54, 61, 91, 94, 95]. En écologie, des dynamiques de populations sont souvent présentées comme des processus de Markov qui correspondent d'un point de vue macroscopique au modèle de Lotka–Volterra cyclique [34, 84]. Finalement, en neurosciences, l'activité qui se propage dans un réseau de neurones est souvent étudiée à partir d'équations différentielles, par exemple par le modèle de Wilson–Cowan [101]. Celui-ci est un autre modèle de champ moyen qui a été reformulé par des processus stochastiques, notamment dans [2, 9, 21].

Une limite importante des modèles de champ moyen est qu'ils négligent d'emblée les corrélations entre les états des unités qu'ils étudient [17, 75, 83]. Ces corrélations pourraient pourtant être importantes, puisque ces modèles sont basés sur des processus stochastiques. Ce problème a suscité beaucoup d'efforts en recherche au cours des dernières décennies, particulièrement en écologie et en épidémiologie. Depuis les années 1990, plusieurs auteurs dont [54, 55, 56, 61, 67, 71, 83, 88, 91, 94] ont présenté des généralisations de modèles épidémiologiques compartimentaux ou de modèles de dynamiques de populations en écologie afin d'inclure explicitement des corrélations dans les équations des modèles. Le même problème survient en neurosciences, où plusieurs auteurs dont [2, 9, 21, 33, 38] ont proposé des méthodes pour permettre ou simuler des corrélations dans des modèles de champ moyen similaires à celui de Wilson–Cowan.

Le travail présenté dans ce mémoire s'inscrit d'abord dans cette lignée. L'objectif premier est de généraliser le modèle de Wilson–Cowan afin d'y inclure explicitement des corrélations. Pour trouver une solution à ce problème, on propose d'abord une nouvelle construction du modèle afin de bien comprendre ses fondations mathématiques. En effet, à la connaissance

de l'auteur de ce mémoire, la littérature ne contient pas de présentation du modèle de Wilson–Cowan où la construction d'un processus de Markov et le passage vers un système d'équations différentielles sont donnés en détail. Lorsque ce modèle est construit, l'attention est portée sur l'interprétation biologique des arguments utilisés bien davantage que sur leur rigueur mathématique. Or, sans bien comprendre la construction mathématique du modèle, il est difficile de bien voir dans quelle mesure il est approximatif et de quelle façon des corrélations entre les états des neurones peuvent avoir un effet.

Par ailleurs, les constructions de plusieurs modèles de champ moyen distincts sont très similaires entre elles, de même que les constructions de leurs versions stochastiques microscopiques. Cependant, du moins à la connaissance de l'auteur de ce mémoire, ils ne sont jamais étudiés de façon aussi générale qu'ils pourraient l'être. Les formes particulières de ces processus sont plutôt étudiées directement dans le contexte d'applications données. Des parallèles sont souvent faits entre des modèles épidémiologiques et écologiques, comme dans [15, 71], mais peu de liens sont faits entre ces modèles et les modèles semblables en neurosciences. Un autre objectif important de ce mémoire est ainsi de présenter la construction d'un modèle suffisamment général pour pouvoir être appliqué à plusieurs phénomènes biologiques distincts, et ce, en faisant le moins de concessions possible quant à la rigueur mathématique des arguments.

Ce mémoire est essentiellement divisé en deux parties : on construit d'abord un modèle théorique, puis on en étudie quelques applications. On commence par présenter au chapitre 1 quelques notions de théorie des probabilités et de théorie des systèmes dynamiques qui seront utilisées dans les chapitres subséquents. Une présentation abstraite des problèmes de fermeture de moments y est aussi donnée. Au chapitre 2, on propose ensuite la construction d'un processus stochastique qui fait évoluer les états des noeuds d'un graphe. On voudra que ce processus soit assez général pour qu'il puisse être appliqué pour représenter différents phénomènes biologiques. On présente également une méthode qui permet d'obtenir un modèle macroscopique de la dynamique moyenne sur le graphe sous la forme d'un système d'équations différentielles ordinaires, de façon à permettre ou non des corrélations entre les états des noeuds. Cette méthode sera basée sur la recherche d'une solution à un problème de moments.

Les deux derniers chapitres sont ensuite consacrés à l'étude d'applications du modèle construit au chapitre 2. D'abord, au chapitre 3, on présente brièvement une application en épidémiologie, à partir de laquelle on retrouvera le modèle classique de Kermack–McKendrick évoqué plus tôt. On présente ensuite différentes méthodes qui ont déjà été proposées dans la littérature et qui permettent d'inclure dans ce modèle des corrélations entre les états des individus. Ce chapitre vise à présenter une application relativement simple des résultats du chapitre 2, qui illustre bien les défis que posent une telle modélisation. Finalement, au chapitre 4, on présente une application du modèle construit au

chapitre 2 en neurosciences. On verra que le modèle obtenu est relié de près au modèle de Wilson–Cowan. On proposera alors deux nouvelles méthodes qui visent à permettre des corrélations entre les états des neurones.

Dans les deux chapitres qui traitent des applications en épidémiologie et en neurosciences, plusieurs résultats numériques sont présentés. Tous ces résultats ont été obtenus dans le langage Python, à l'aide du paquetage PopNet [77], écrit par l'auteur de ce mémoire. Quelques commentaires sur ce paquetage et sur la méthodologie utilisée pour obtenir les résultats numériques sont donnés à l'annexe B. On peut aussi noter que les figures 2.1, 3.1, 3.3 et 4.1 ont été réalisées directement à l'aide du paquetage TikZ en L^AT_EX, et que toutes les autres ont plutôt été réalisées dans le langage Python, à l'aide de la librairie Matplotlib [53].

Chapitre 1

Systèmes dynamiques et probabilités

Ce chapitre présente une revue de quelques notions importantes de systèmes dynamiques et de théorie des probabilités qui seront utilisées dans les chapitres subséquents.

Les deux premières sections traitent respectivement de systèmes dynamiques et de théorie des probabilités. Les notions de théorie des probabilités seront utilisées tout au long du mémoire, mais particulièrement au chapitre 2 dans la construction d'un processus stochastique qui fait évoluer les états des noeuds d'un graphe. Les notions de systèmes dynamiques, elles, seront principalement utilisées lors de l'analyse des systèmes dynamiques obtenus dans les applications des résultats du chapitre 2 en biologie mathématique présentées aux chapitres 3 et 4. Finalement, on présente dans la troisième section une brève introduction aux problèmes de fermeture de moments, où l'on s'intéresse à l'approximation d'équations différentielles qui font évoluer les moments de variables aléatoires. On verra qu'un tel problème survient naturellement dans la recherche d'un modèle macroscopique à partir du processus stochastique présenté au chapitre 2.

§ 1.1. Notions de théorie des systèmes dynamiques

On introduit dans cette section quelques notions de théorie des systèmes dynamiques. On suppose que le lecteur s'est déjà familiarisé avec l'étude des équations différentielles linéaires ; pour ces notions on réfère au chapitre 1 de [80]. La présentation donnée ici est principalement basée sur les livres de Perko [80] et de Schaeffer et Cain [89]. Par souci de concision, on omettra les preuves des résultats, qui sont tous classiques.

§ 1.1.1. Équations différentielles ordinaires

On s'intéresse ici à des équations différentielles ordinaires d'ordre 1, c'est-à-dire à des équations de la forme

$$\dot{x} = f(x, t), \quad (1.1)$$

où $f: E \times I \rightarrow \mathbb{R}^n$ pour un ouvert E de \mathbb{R}^n et un intervalle $I \subset \mathbb{R}$, et où le point indique une dérivée par rapport au paramètre t . Dans les applications, il arrive souvent qu'il soit plus pratique d'écrire une telle équation en explicitant chacune des composantes. Dans ce

cas, on dit parfois qu'il s'agit d'un *système* d'équations différentielles ordinaires. D'un point de vue théorique, cependant, ces deux concepts sont identiques.

On étudiera ici principalement des équations différentielles autonomes.

Définition 1.1. Une équation différentielle de la forme (1.1) est dite *autonome* lorsque f ne dépend pas de t . Autrement, on dit qu'elle est *non autonome*. \diamond

Pour la suite, on se concentrera majoritairement sur des équations différentielles autonomes, qui ont alors la forme

$$\dot{x} = f(x). \quad (1.2)$$

Pour le reste de la section, on notera toujours $f: E \rightarrow \mathbb{R}^n$ pour un ouvert E de \mathbb{R}^n .

La première question d'importance à l'étude d'une telle équation est l'existence et l'unicité de ses solutions. On dit qu'une fonction x est une solution de l'équation différentielle (1.2) sur un intervalle $I \subset \mathbb{R}$ si x est dérivable sur I et si pour tout $t \in I$, $\dot{x}(t) = f(x(t))$. En général, une seule équation différentielle a cependant de nombreuses solutions : celles-ci dépendent en effet d'un point de départ choisi dans E . On étudie donc plutôt, pour un $x_0 \in E$ donné, le problème aux conditions initiales

$$\dot{x} = f(x), \quad x(0) = x_0. \quad (1.3)$$

Sans surprise, l'existence des solutions d'un tel problème dépend des propriétés de f . La propriété suivante est suffisante.

Définition 1.2. On dit que f est *localement Lipschitz* sur E si pour tout $x \in E$, il existe un voisinage ouvert U de x et un $L \geq 0$ tels que $y, z \in U$, $|f(y) - f(z)| \leq L|y - z|$. \diamond

Le résultat suivant donne un critère utile pour déterminer si une fonction a cette propriété.

Proposition 1.1. Si $f \in \mathcal{C}^1(E)$, alors f est localement Lipschitz sur E . \blacklozenge

On peut maintenant donner le théorème classique d'existence et d'unicité des solutions.

Théorème 1.2. Soit $x_0 \in E$. Si f est localement Lipschitz sur E , alors il existe un intervalle $(-\varepsilon, \varepsilon)$ et une fonction $x: (-\varepsilon, \varepsilon) \rightarrow E$ de classe \mathcal{C}^1 qui est une solution du problème (1.3). De plus, cette solution est unique, au sens où si $y: (-\eta, \eta) \rightarrow E$ est une autre solution du même problème, alors $y(t) = x(t)$ pour tout $t \in (-\varepsilon, \varepsilon) \cap (-\eta, \eta)$. \blacklozenge

Il existe même un intervalle maximal sur lequel la solution du problème (1.3) existe.

Théorème 1.3. Si f est localement Lipschitz sur E , alors pour chaque $x_0 \in E$, il existe un intervalle I et une solution $x: I \rightarrow E$ du problème (1.3) tels que si $y: J \rightarrow E$ est une autre solution du problème (1.3), alors $J \subset I$ et $y(t) = x(t)$ pour tout $t \in J$. \blacklozenge

Remarque. L'intervalle I du théorème précédent est souvent appelé l'*intervalle maximal d'existence* du problème (1.3). Cet intervalle dépend de la condition initiale x_0 , il est donc souvent noté $I(x_0)$.

Pour étudier simultanément les solutions d'une équation différentielle pour différentes conditions initiales, il est utile d'introduire la notion de flot.

Définition 1.3. Si f est localement Lipschitz sur E , pour chaque $x_0 \in E$, soit $\phi(x_0, t)$ la solution du problème (1.3) définie sur son intervalle maximal d'existence $I(x_0)$. Alors pour $t \in I(x_0)$, les applications $\phi_t: x_0 \mapsto \phi(x_0, t)$ sont appelées le *flot* de l'équation différentielle (1.2) ou du champ vectoriel f . \diamond

Certains ensembles ne sont pas modifiés par le flot d'une équation différentielle.

Définition 1.4. Soit ϕ_t le flot de l'équation différentielle (1.2). On dit qu'un ensemble $S \subset E$ est *invariant par le flot* ϕ_t si $\phi_t(x_0) \in S$ pour tout $x_0 \in E$ et pour tout $t \in I(x_0)$. On dit que S est *positivement* ou *négativement* invariant par le flot ϕ_t si $\phi_t(x_0) \in S$ pour tout $x_0 \in E$ et pour tout $t \in I(x_0)$ tel que $t \geq 0$ ou tel que $t \leq 0$. \diamond

Autrement dit, un ensemble est invariant par le flot d'une équation différentielle si toutes les solutions de l'équation différentielle dont la condition initiale s'y trouve ont toujours été et seront toujours dans cet ensemble. Les conditions d'être positivement ou négativement invariant sont moins fortes : un ensemble est positivement invariant si les solutions de l'équation différentielle n'en sortent jamais une fois qu'elles y sont entrées, et il est négativement invariant si les solutions qui s'y trouvent à un instant donné y ont toujours été.

En pratique, il est souvent possible d'étudier de façon géométrique si un ensemble est invariant ou non, en étudiant l'orientation du champ vectoriel f associé à l'équation différentielle. Le théorème suivant est un résultat très utile en ce sens, souvent appelé le théorème de Bony–Brezis en référence aux travaux de Bony [6] et de Brezis [11], quoique initialement découvert par Nagumo [73]. L'énoncé qu'on donne ici est basé sur l'énoncé donné dans [50].

Théorème 1.4. Soit F un fermé de E . Si f est localement Lipschitz sur E , alors les conditions suivantes sont équivalentes :

- i) Chaque solution de (1.3) avec $x_0 \in F$ est entièrement contenue dans F ;
- ii) $\langle f(x), n_x \rangle \leq 0$ pour tout $x \in F$ et tout n_x normal à F en x vers l'extérieur.

◆

Certains ensembles invariants « attirent » les solutions de l'équation différentielle.

Définition 1.5. Soit $A \subset E$ un fermé invariant par le flot de (1.2). On suppose que pour tout $t \in \mathbb{R}$, le flot $\phi_t: E \rightarrow E$ est défini. On dit que A est un *ensemble attractif* de (1.2) s'il existe un voisinage ouvert U de A tel que pour tout $x \in U$, $\phi_t(x) \in U$ pour tout $t \geq 0$ et $d(\phi_t(x), A) \rightarrow 0$ lorsque $t \rightarrow \infty$, où d note la distance euclidienne. \diamond

Remarque. La réunion de tous les voisinages U qu'on pourrait choisir dans la définition précédente est parfois appelée le *bassin d'attraction* d'un ensemble attractif A .

§ 1.1.2. Solutions fermées

Les solutions d'une équation différentielle sont souvent organisées autour d'un petit nombre de solutions fermées vers lesquelles toutes les solutions convergent. On introduit ici quelques cas simples de telles solutions.

Définition 1.6. On dit qu'un point $x \in E$ est un *point fixe* de l'équation différentielle (1.2) si $f(x) = x$. Un point fixe x est dit *hyperbolique* si les parties réelles des valeurs propres de $Df(x)$ sont toutes non nulles. Il est dit *stable* si pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que pour tout $y \in E$ avec $|y - x| < \delta$ et pour tout $t \geq 0$, $|\phi_t(y) - x| < \varepsilon$. Sinon, il est dit *instable*. \diamond

Il est bon de rappeler que pour les systèmes linéaires, la stabilité d'un point fixe hyperbolique est déterminée par les valeurs propres de la matrice jacobienne. En effet, dans ce cas, le point fixe est stable si les parties réelles des valeurs propres sont toutes négatives. Il est possible de transférer ce résultat aux systèmes non linéaires à partir du théorème suivant, démontré indépendamment par Hartman dans [45] et par Grobman dans [41].

Théorème 1.5. Soit ϕ_t le flot de (1.2), où f est supposée de classe \mathcal{C}^1 . Si l'origine est un point fixe hyperbolique de (1.2), alors il existe un homéomorphisme $H: U \rightarrow V$, où U et V sont deux ouverts de \mathbb{R}^n qui contiennent l'origine, tel que pour chaque $x \in U$, il existe un intervalle $I \subset \mathbb{R}$ qui contient zéro tel que pour tout $t \in I$, $H \circ \phi_t(x) = e^{Df(0)t} H(x)$. \blacklozenge

Remarque. Si $A \in \mathbb{R}^{n \times n}$, la solution de l'équation différentielle linéaire $\dot{x} = Ax$ avec la condition initiale $x(0) = x_0$ est donnée par $x(t) = e^{At}x_0$, où l'exponentielle de la matrice At est définie par sa série de Taylor

$$e^{At} := \sum_{n=0}^{\infty} \frac{A^n t^n}{n!}, \quad (1.4)$$

qui converge absolument pour tout $t \in \mathbb{R}$. Ainsi, le théorème de Hartman–Grobman montre une correspondance entre les trajectoires du système non linéaire (1.2) et celles du système linéaire $\dot{x} = Df(0)x$. Ceci montre que dans le cas d'un point fixe hyperbolique, la stabilité est déterminée par les parties réelles des valeurs propres de la matrice jacobienne.

Les cycles sont une autre forme importante de solutions fermées.

Définition 1.7. Un *cycle* de l'équation différentielle (1.2) est une solution de (1.2) qui est une courbe fermée mais qui n'est pas un point fixe. Un cycle Γ est dit *stable* si pour tout $\varepsilon > 0$ il existe un voisinage U de Γ tel que si $x \in U$ et $t \geq 0$, alors $d(\phi_t(x), \Gamma) < \varepsilon$, où ϕ_t note le flot de f et d note la distance euclidienne. Un cycle est dit *instable* lorsqu'il n'est pas stable. \diamond

L'exemple suivant donne une bonne illustration de la présence de points fixes et de cycles. On en profite pour introduire la notion de portrait de phase, un outil fort utile pour visualiser le comportement d'un système.

Exemple 1.1 – Le modèle de Lotka–Volterra. Le modèle de Lotka–Volterra, introduit par Lotka [69] et par Volterra [98], est un modèle écologique qui décrit comment varient les populations d'un prédateur et d'une proie. Si x et y notent respectivement les quantités de proies et de prédateurs, le modèle propose que

$$\dot{x} = \alpha x - \beta xy, \quad (1.5a)$$

$$\dot{y} = \gamma xy - \delta y, \quad (1.5b)$$

où α , β , γ et δ sont des paramètres strictement positifs qui décrivent les interactions entre les deux espèces.

Il est facile de vérifier que ce système comporte deux points fixes. D'abord, il est clair que $\dot{x} = \dot{y} = 0$ lorsque $x = y = 0$. L'origine est donc un point fixe de (1.5). On remarque d'ailleurs que si l'une des coordonnées est nulle, la seule possibilité pour que $\dot{x} = \dot{y} = 0$ est que les deux coordonnées soient nulles. Ensuite, on voit que $\dot{x} = 0$ lorsque $y = \alpha/\beta$ et que $\dot{y} = 0$ lorsque $x = \delta/\gamma$. Le point $(x, y) = (\delta/\gamma, \alpha/\beta)$ est donc un autre point fixe de (1.5).

Soit $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ le champ vectoriel associé au système (1.5). Pour étudier la stabilité de ces points fixes, on peut y évaluer la matrice jacobienne. Des calculs directs montrent que

$$Df(0,0) = \begin{pmatrix} \alpha & 0 \\ 0 & -\delta \end{pmatrix} \quad \text{et que} \quad Df(\delta/\gamma, \alpha/\beta) = \begin{pmatrix} 0 & -\beta\delta/\gamma \\ \alpha\gamma/\beta & 0 \end{pmatrix}. \quad (1.6)$$

Comme $\alpha > 0$ et $\delta > 0$, on voit que $Df(0,0)$ a deux valeurs propres réelles non nulles, l'une positive et l'autre négative. Le théorème d'Hartman–Grobman implique alors que ce point fixe est instable. Cependant, on peut vérifier que les valeurs propres de $Df(\delta/\gamma, \alpha/\beta)$ sont $\pm i\sqrt{\alpha\delta}$. Comme leurs parties réelles sont nulles, le théorème d'Hartman–Grobman ne permet pas de tirer de conclusion quant à leur stabilité.

Afin d'observer ce qui se produit autour du point fixe $(\delta/\gamma, \alpha/\beta)$, il est instructif de tracer le portrait de phase du système, c'est-à-dire de représenter graphiquement le champ vectoriel associé à l'équation différentielle. Celui-ci est présenté à la figure 1.1. On y constate qu'en fait, les solutions autour du point fixe $(\delta/\gamma, \alpha/\beta)$ sont des cycles !

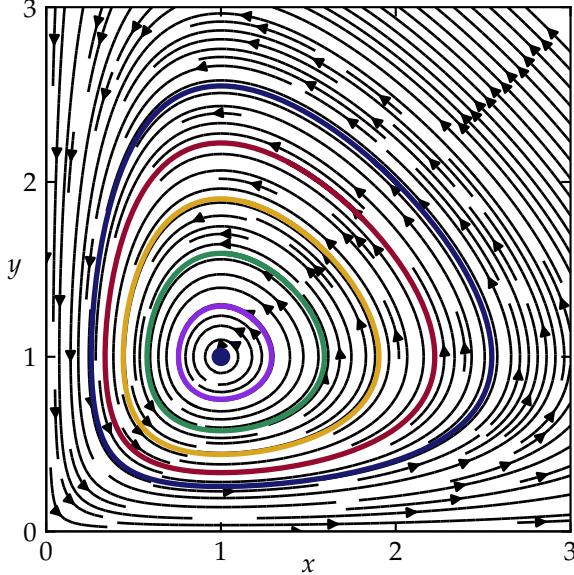


FIGURE 1.1 – Portrait de phase du système de Lotka–Volterra (1.5) pour $\alpha = \beta = \gamma = \delta = 1$. Chacune des courbes correspond à une solution du système pour une différente condition initiale, et le point bleu est le point fixe $(x, y) = (\delta/\gamma, \alpha/\beta)$.

Contrairement aux cycles présents dans l'exemple précédent, un cycle peut aussi être une solution vers laquelle d'autres solutions convergent. Pour clarifier ce cas, on introduit les notions suivantes.

Définition 1.8. Soit $x \in E$. Un point $p \in E$ est un *point ω -limite* (resp. *point α -limite*) d'une solution $t \mapsto \phi(x, t)$ de (1.2) s'il existe une suite $(t_n)_{n=1}^{\infty}$ avec $t_n \rightarrow \infty$ (resp. $t_n \rightarrow -\infty$) telle que $\phi(x, t_n) \rightarrow p$ lorsque $n \rightarrow \infty$. Les ensembles des points ω -limites et α -limites d'une même solution sont appelés l'*ensemble ω -limite* et l'*ensemble α -limite* de cette solution. ◇

On peut alors définir un cycle limite dans \mathbb{R}^2 .

Définition 1.9. Si $E \subset \mathbb{R}^2$ et $f: E \rightarrow \mathbb{R}^2$, un *cycle limite* de (1.2) est un cycle Γ de (1.2) qui est aussi un ensemble α - ou ω -limite d'une solution de (1.2) différente de Γ . Dans ce cas, si Γ est l'ensemble ω -limite (resp. α -limite) de toutes les solutions dans un voisinage de Γ , alors Γ est un cycle limite *stable* (resp. *instable*). ◇

§ 1.1.3. Bifurcations de Hopf

Il arrive fréquemment que différentes propriétés importantes d'une équation différentielle, comme la position ou la stabilité de ses points fixes, dépendent des valeurs de certains paramètres. Par exemple, on a vu à l'exemple 1.1 que la position de l'un des points fixes du système de Lotka–Volterra dépend des paramètres α, β, γ et δ . Ainsi, lorsque la valeur d'un paramètre d'un système varie, son portrait de phase peut changer drastiquement ; dans ce cas on dit qu'une *bifurcation* se produit.

La théorie des bifurcations s'avère compliquée, et une bonne introduction à ses concepts fondamentaux nécessiterait d'introduire de nombreux concepts préalables et sortirait du cadre de ce mémoire. Pour une bonne introduction à cette théorie, on réfère au chapitre 4 du livre de Perko [80], et pour une présentation plus complète on réfère à l'ouvrage classique de Guckenheimer et Holmes [43].

Dans ce mémoire, on ne s'intéressera qu'à un type particulier de bifurcations : les *bifurcations de Hopf* (aussi parfois appelées *bifurcations de Poincaré–Andronov–Hopf* ; voir la discussion historique à la fin de la section 20.2 du livre de Wiggins [100]). Plutôt que de les introduire rigoureusement dans toute leur généralité, on ne présente ici qu'un exemple qui en illustre bien les idées principales.

On considère le système

$$\dot{x} = \beta x - y - \alpha(x^2 + y^2)x, \quad (1.7a)$$

$$\dot{y} = x + \beta y - \alpha(x^2 + y^2)y \quad (1.7b)$$

où $\beta \in \mathbb{R}$ et $\alpha \neq 0$. Il est facile de vérifier que ce système a un unique point fixe situé à

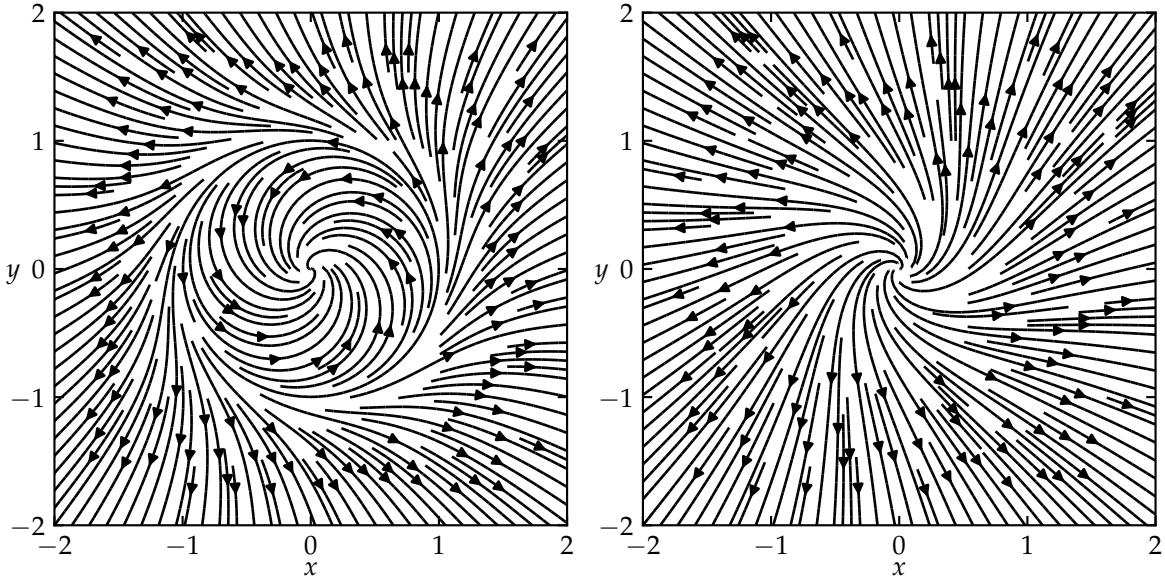


FIGURE 1.2 – Portraits de phase du système (1.7) si $\alpha = -1$, avec $\beta = -1$ à gauche et $\beta = 1$ à droite.

l'origine. Soit $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ le champ vectoriel associé à (1.7). Un calcul direct montre que

$$Df(0,0) = \begin{pmatrix} \beta & -1 \\ 1 & \beta \end{pmatrix}. \quad (1.8)$$

Les valeurs propres de cette matrice sont $\beta \pm i$. Ainsi, le théorème d'Hartman–Grobman permet de conclure que le point fixe à l'origine est stable lorsque $\beta < 0$, et instable lorsque $\beta > 0$. Lorsque β change de signe, la partie réelle d'une paire de valeurs propres conjuguées de Df change de signe. C'est lorsque ce phénomène survient qu'on dit qu'une bifurcation de Hopf se produit.

Pour mieux comprendre le phénomène, il est instructif d'étudier la coordonnée polaire $r := \sqrt{x^2 + y^2}$. En effet, à partir de la relation $r\dot{r} = x\dot{x} + y\dot{y}$, on vérifie aisément que

$$r\dot{r} = \beta r^2 - \alpha r^4, \quad \text{donc que} \quad \dot{r} = \alpha r(\beta/\alpha - r^2). \quad (1.9)$$

Ainsi, si β et α sont de même signe, le cercle de rayon $\sqrt{\beta/\alpha}$ est invariant par le flot du système (1.7).

Deux cas sont ensuite à considérer séparément. D'abord, on suppose que $\alpha < 0$. Si $\beta < 0$, alors $\dot{r} \geq 0$ lorsque $r \geq \sqrt{\beta/\alpha}$, et si $\beta > 0$, alors $\dot{r} > 0$ pour tout $r > 0$. Dans ce cas, on voit donc que pour $\beta < 0$ il existe un cycle limite instable de rayon $\sqrt{\beta/\alpha}$, alors que pour $\beta > 0$ ce cycle limite est disparu et il ne reste que l'unique point fixe instable à l'origine. Deux portraits de phase pour ce système sont illustrés à la figure 1.2.

Ensuite, on suppose que $\alpha > 0$. Si $\beta < 0$, alors $\dot{r} < 0$, et si $\beta > 0$, alors $\dot{r} \geq 0$ lorsque $r \leq \sqrt{\beta/\alpha}$. Dans ce cas, on voit donc que pour $\beta < 0$ il n'y a qu'un point fixe stable à l'origine, mais

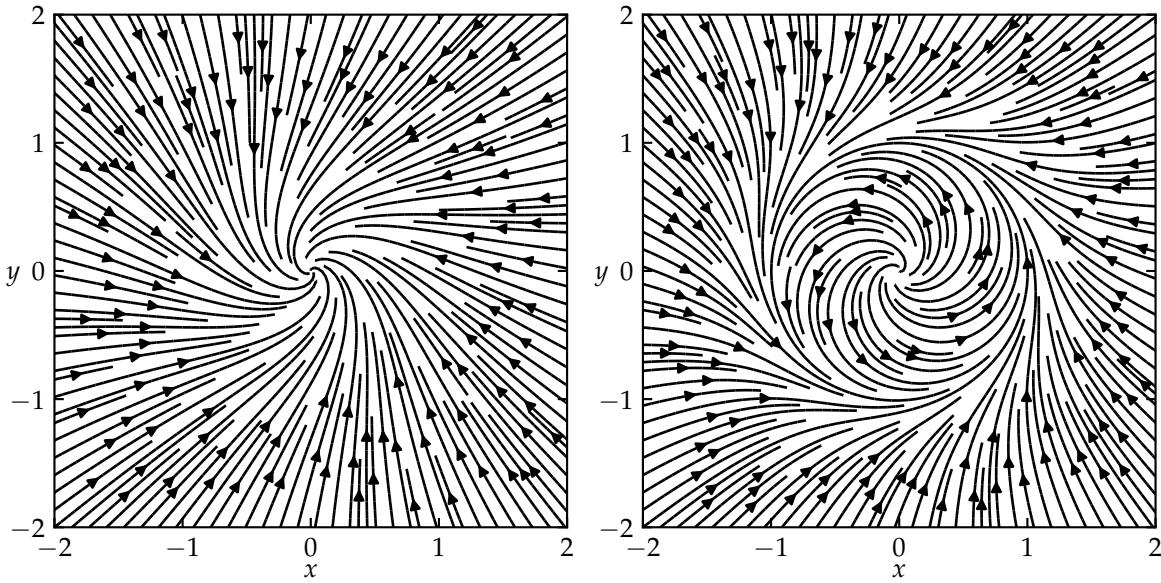


FIGURE 1.3 – Portraits de phase du système (1.7) si $\alpha = 1$, avec $\beta = -1$ à gauche et $\beta = 1$ à droite.

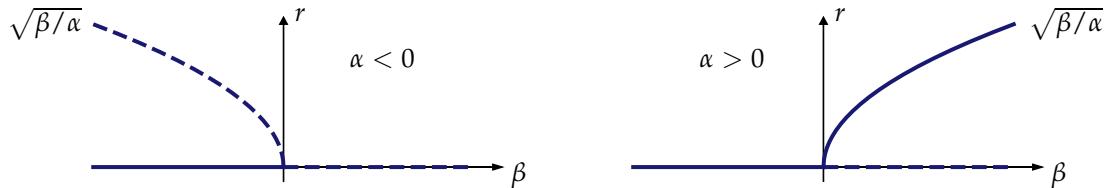


FIGURE 1.4 – Diagrammes de bifurcation pour le système (1.7) pour la bifurcation de Hopf sous-critique ($\alpha < 0$) et supercritique ($\alpha > 0$) en fonction du paramètre β . Sur les deux diagrammes, la courbe $\sqrt{\beta/\alpha}$ indique le rayon du cycle limite, et la droite superposée à l'axe horizontal indique le point fixe à l'origine. Dans les deux cas, une ligne pleine indique un état stable et une ligne pointillée indique un état instable.

que lorsque $\beta > 0$, le point fixe devient instable et un cycle limite stable de rayon $\sqrt{\beta/\alpha}$ apparaît. Deux portraits de phase pour ce système sont illustrés à la figure 1.3.

Ces deux comportements sont caractéristiques de deux types distincts de bifurcations de Hopf. Dans le cas où $\alpha < 0$, lorsque le paramètre β passe de négatif à positif, on observe la disparition d'un cycle limite instable ; la bifurcation de Hopf est alors dite *sous-critique*. Dans le cas où $\alpha > 0$, lorsque le paramètre β passe de négatif à positif, on observe plutôt l'apparition d'un cycle limite stable ; la bifurcation de Hopf est alors dite *supercritique*.

Ces bifurcations peuvent être visualisées graphiquement en traçant un *diagramme de bifurcation*, qui représente les points fixes et les cycles limites du système en fonction du paramètre par rapport auquel une bifurcation se produit, qui est β dans ce cas-ci. De tels diagrammes sont présentés à la figure 1.4 pour le cas où $\alpha < 0$ et pour le cas où $\alpha > 0$.

§ 1.2. Notions de théorie des probabilités

Cette section donne une brève introduction à certains concepts de la théorie des probabilités. On suppose ici que le lecteur connaît déjà les rudiments de la théorie de la mesure; pour ces notions on réfère aux ouvrages classiques de Halmos [44] et de Rudin [86].

L'essentiel du matériel présenté dans cette section tire son origine des travaux de Kolmogorov publiés pour la première fois dans [64] en 1933 (voir [63] pour une traduction anglaise). Ces résultats sont aujourd'hui standards et sont discutés dans la plupart des livres qui traitent de la théorie des probabilités. La présentation donnée ici est majoritairement basée sur le livre de Çınlar [18]. Par souci de concision, on omettra les preuves des résultats qui sont tirés directement de [18] ou d'autres références citées dans le texte.

§ 1.2.1. Espaces de probabilité et variables aléatoires

Sans surprise, la première structure d'importance dans l'étude de la théorie des probabilités est l'espace de probabilité.

Définition 1.10. Un *espace de probabilité* est un espace mesuré $(\Omega, \mathcal{F}, \mathbb{P})$ tel que $\mathbb{P}(\Omega) = 1$. Dans ce cas, on dit que \mathbb{P} est une *mesure de probabilité*. De plus, on appelle les éléments de \mathcal{F} des *événements*, et pour $E \in \mathcal{F}$ on appelle $\mathbb{P}(E)$ la *probabilité* de E . ◇

Exemple 1.2. L'exemple par excellence de l'espace de probabilité est celui qui sert à représenter un lancer de dé. Dans ce cas, on prend $\Omega := \{1, 2, 3, 4, 5, 6\}$ et $\mathcal{F} := 2^\Omega$, l'ensemble puissance de Ω . On définit ensuite simplement la mesure \mathbb{P} sur (Ω, \mathcal{F}) par $\mathbb{P}(E) := |E|/6$, où $|E|$ note le cardinal de E . On a alors bien que $\mathbb{P}(\Omega) = 1$, donc l'espace mesuré $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace de probabilité. ▀

Exemple 1.3. Soient $\mathcal{B}([0, 1])$ la σ -algèbre de Borel sur $[0, 1]$ et λ la mesure de Lebesgue sur $[0, 1]$. Puisque $\lambda([0, 1]) = 1$, alors l'espace mesuré $([0, 1], \mathcal{B}([0, 1]), \lambda)$ est un espace de probabilité. ▀

Pour le reste de la sous-section, on convient que $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace de probabilité. Afin de représenter des expériences aléatoires sur cet espace, on utilise des fonctions sur Ω .

Définition 1.11. Soient un espace mesurable (Γ, \mathcal{G}) et une fonction $X: \Omega \rightarrow \Gamma$. Si X est mesurable, alors on dit que X est une *variable aléatoire à valeurs dans Γ* . Dans ce cas, la mesure-image $\mathbb{P}_X: \mathcal{G} \rightarrow [0, 1]$ de \mathbb{P} par X , qui est définie par $\mathbb{P}_X(E) := \mathbb{P}(X^{-1}(E))$, est appelée la *loi* de X . Lorsque des variables aléatoires suivent la même loi, on dit qu'elles sont *identiquement distribuées*. ◇

Remarque. Dans le cas d'une variable aléatoire X à valeurs dans \mathbb{R} , si la loi de X est absolument continue par rapport à la mesure de Lebesgue, alors le théorème de Radon–Nikodym prévoit l'existence d'une fonction borélienne $p: \mathbb{R} \rightarrow [0, \infty]$ telle que $\mathbb{P}_X(B) = \int_B p \, d\lambda$ pour

chaque $B \in \mathcal{B}(\mathbb{R})$. Dans ce cas, on dit que X est à *densité*, et on appelle p la *densité* de X . On remarque que le théorème de Radon–Nikodym garantit également que p est unique p.s.

Plus loin, on s'intéressera particulièrement à deux caractéristiques particulières que peuvent posséder des lois de probabilité sur \mathbb{R} qui possèdent une densité.

Définition 1.12. Une loi de probabilité sur \mathbb{R} de densité f est dite *unimodale* si pour un certain $a \in \mathbb{R}$, f est croissante sur $(-\infty, a)$, décroissante sur (a, ∞) , et a un maximum unique en a . \diamond

Définition 1.13. Une loi de probabilité sur \mathbb{R} de densité f est dite *symétrique* autour d'un $a \in \mathbb{R}$ si $f(a+x) = f(a-x)$ pour tout $x \in \mathbb{R}$. \diamond

Remarque. Si une loi de probabilité est unimodale et symétrique autour d'un point a , alors sa densité doit être maximale en a .

L'intégrale d'une variable aléatoire a un sens très important en théorie des probabilités.

Définition 1.14. Soit X une variable aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$. L'*espérance* de X , notée $\mathbb{E}[X]$, est son intégrale sur Ω par rapport à la mesure \mathbb{P} . \diamond

Remarque. Comme pour l'intégrale d'une fonction mesurable en général, il peut arriver que l'espérance d'une variable aléatoire n'existe pas.

L'espérance d'une variable aléatoire X peut être interprétée comme la valeur attendue en moyenne si l'on répète l'expérience associée à X un très grand nombre de fois. Cette intuition provient de la loi des grands nombres, qui est présentée plus loin.

Le lien entre la probabilité et l'espérance devient très intuitif pour une variable aléatoire simple à valeurs réelles. Par exemple, étant données une partition $\{A_j\}_{j=1}^n \subset \mathcal{F}$ de Ω et des constantes $c_1, \dots, c_n \in \mathbb{R}$, soit $X := \sum_{j=1}^n c_j \mathbb{1}_{A_j}$ où $\mathbb{1}_{A_j}$ est la fonction indicatrice de l'ensemble A_j . Alors X est une variable aléatoire sur Ω à valeurs dans \mathbb{R} , et

$$\mathbb{E}[X] = \int_{\Omega} X \, d\mathbb{P} = \int_{\Omega} \sum_{j=1}^n c_j \mathbb{1}_{A_j} \, d\mathbb{P} = \sum_{j=1}^n c_j \mathbb{P}(A_j). \quad (1.10)$$

Ainsi, l'espérance de X correspond à la somme de chaque valeur que peut prendre X pondérée par la probabilité que X prenne cette valeur. Cette intuition peut se généraliser à n'importe quelle variable aléatoire à valeurs réelles, à partir du fait que chaque fonction positive mesurable est la limite d'une suite de fonctions simples.

On peut se demander à quel point une variable aléatoire peut prendre des valeurs éloignées de son espérance. La variance permet de quantifier cette dispersion.

Définition 1.15. Soit $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. L'expression $\text{Var}[X] := \mathbb{E}[|X - \mathbb{E}[X]|^2]$ est appelée la *variance* de X . \diamond

Remarque. $\text{Var}[X] = \|X - \mathbb{E}[X]\|_2^2$, où $\|\cdot\|_2$ est la norme sur $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Il peut être pratique de pouvoir borner la variance d'une variable aléatoire. Dans [81], Popoviciu a donné une borne pour le cas de variables aléatoires à valeurs réelles bornées.

Théorème 1.6 – Inégalité de Popoviciu. Soit $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Si $a \leq X \leq b$ p.s. pour certains $a, b \in \mathbb{R}$, alors $\text{Var}[X] \leq \frac{1}{4}(b - a)^2$. ◆

Il est possible de donner une borne supérieure plus stricte que celle donnée par l'inégalité de Popoviciu en faisant intervenir l'espérance de la variable aléatoire d'intérêt. Le théorème suivant, prouvé par Bhatia et Davis dans [4], fournit une telle borne.

Théorème 1.7 – Inégalité de Bhatia–Davis. Soit $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Si $a \leq X \leq b$ p.s. pour certains $a, b \in \mathbb{R}$, alors $\text{Var}[X] \leq (b - \mathbb{E}[X])(\mathbb{E}[X] - a)$. ◆

Remarque. Cette inégalité est optimale. En effet, les bornes sont atteintes pour X binaire avec $\mathbb{P}[X = b] = p$ et $\mathbb{P}[X = a] = 1 - p$ pour un $p \in [0, 1]$. Dans ce cas, $\mathbb{E}[X] = bp + a(1 - p)$, de sorte que $b - \mathbb{E}[X] = (b - a)(1 - p)$ et que $\mathbb{E}[X] - a = (b - a)p$. Or, $\mathbb{E}[X^2] = b^2p + a^2(1 - p)$, donc la variance de X est justement

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = a^2 + p(b^2 - a^2) - (a + p(b - a))^2 \quad (1.11)$$

$$= p(b^2 - a^2) - 2ap(b - a) - p^2(b - a)^2 = p(1 - p)(b - a)^2. \quad (1.12)$$

On conclut cette sous-section en introduisant la fonction de répartition d'une variable aléatoire. En probabilités élémentaires, les fonctions de répartition sont souvent utilisées comme des alternatives aux lois de variables aléatoires à valeurs réelles.

Définition 1.16. Soit X une variable aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans \mathbb{R} . La fonction $F_X : \mathbb{R} \rightarrow [0, 1]$ définie par $F_X(x) := \mathbb{P}[X \leq x]$ est appelée la *fonction de répartition* de X . ◇

Remarque. On peut montrer facilement à partir des propriétés des mesures qu'une fonction de répartition est toujours croissante et continue à droite.

Les fonctions de répartition de lois de probabilités unimodales et symétriques ont certaines propriétés intéressantes, qui seront utiles pour la suite.

Proposition 1.8. Soit $F : \mathbb{R} \rightarrow [0, 1]$ la fonction de répartition d'une loi de probabilité unimodale et symétrique autour de $a \in \mathbb{R}$, qui possède une densité. Si F est de classe \mathcal{C}^k pour un $k \in \mathbb{N}$, alors

i) Si $k \geq 2$, alors F'' est positive sur $(-\infty, a)$, nulle en a , et négative sur (a, ∞) .

De plus, pour $\ell \in \mathbb{N}$ avec $1 \leq \ell \leq k$:

ii) Si ℓ est impair, $F^{(\ell)}$ est paire autour de a , c.-à-d. que $\forall x \in \mathbb{R}, F^{(\ell)}(a + x) = F^{(\ell)}(a - x)$;

iii) Si ℓ est pair, $F^{(\ell)}$ est impaire autour de a , c.-à-d. que $\forall x \in \mathbb{R}, F^{(\ell)}(a + x) = -F^{(\ell)}(a - x)$. ◆

Démonstration. Puisque la loi de probabilité est unimodale avec son maximum en a , F' , qui correspond à sa densité, est croissante sur $(-\infty, a)$ et décroissante sur (a, ∞) . Si $k \geq 2$, alors F'' existe et est continue partout. Dans ce cas, elle est positive sur $(-\infty, a)$ et négative sur (a, ∞) . Comme elle est continue, alors $F''(a) = 0$. Ceci termine de montrer i).

Les résultats *ii)* et *iii)* peuvent être prouvés du même coup, en montrant que pour tout ℓ entre 1 et k , $F^{(\ell)}(a+x) = (-1)^{(1+\ell)}F^{(\ell)}(a-x)$. Ce résultat est vrai pour $\ell = 1$, car F' est la densité de la loi de probabilité et que celle-ci est symétrique. On suppose que le résultat tienne pour un $\ell \leq k-1$. On sait alors que la dérivée $(\ell+1)$ -ième de F existe, donc on peut directement dériver chaque côté de la relation précédente pour obtenir que

$$F^{(\ell+1)}(a+x) = (-1)^{2+\ell}F^{(\ell+1)}(a-x) = (-1)^\ell F^{(\ell+1)}(a-x). \quad (1.13)$$

Par induction, les résultats *ii)* et *iii)* sont donc vrais pour tout ℓ entre 1 et k . ■

§ 1.2.2. Indépendance et espérance conditionnelle

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Étant données deux variables aléatoires X et Y sur cet espace, on peut se demander s'il existe une relation entre elles. Lorsque l'on réalise l'expérience associée à X , a-t-on davantage d'information sur l'éventuel résultat de l'expérience associée à Y ? Cette question est d'abord reliée à une possible dépendance entre X et Y , laquelle est reliée aux σ -algèbres engendrées par ces variables aléatoires.

Définition 1.17. Soit X une variable aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans un espace mesurable (Γ, \mathcal{G}) . La σ -algèbre engendrée par X est l'ensemble $\sigma(X) := \{X^{-1}(A) : A \in \mathcal{G}\}$. ◊

Remarque 1. Il est facile de vérifier à partir des propriétés des préimages que $\sigma(X)$ est effectivement une σ -algèbre. Ainsi, $\sigma(X)$ est la plus petite σ -algèbre qui rende X mesurable.

Remarque 2. Puisque par définition d'une variable aléatoire X doit être \mathcal{F} -mesurable, $\sigma(X)$ est nécessairement une sous- σ -algèbre de \mathcal{F} , c'est-à-dire que $\sigma(X) \subset \mathcal{F}$.

Définition 1.18. Soient \mathcal{G} et \mathcal{H} deux sous- σ -algèbres de \mathcal{F} . On dit que \mathcal{G} et \mathcal{H} sont *indépendantes* si pour tout $A \in \mathcal{G}$ et pour tout $B \in \mathcal{H}$, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. De même, on dit que deux variables aléatoires X et Y sur $(\Omega, \mathcal{F}, \mathbb{P})$ sont *indépendantes* si $\sigma(X)$ et $\sigma(Y)$ sont indépendantes. ◊

Le théorème suivant est un résultat important sur les variables aléatoires indépendantes.

Théorème 1.9. Soient X et Y deux variables aléatoires indépendantes sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans les espaces mesurables (Γ, \mathcal{G}) et dans (Θ, \mathcal{H}) respectivement. Soient $f: \Gamma \rightarrow \mathbb{R}$ et $g: \Theta \rightarrow \mathbb{R}$ deux fonctions mesurables. Alors $\mathbb{E}[(f \circ X)(g \circ Y)] = \mathbb{E}[f \circ X]\mathbb{E}[g \circ Y]$. ◆

Si des variables aléatoires sont dépendantes, il peut être utile de pouvoir quantifier « à quel point » elles sont dépendantes. Une façon de le faire est d'étudier la relation linéaire entre elles, qui est quantifiée par leur covariance.

Définition 1.19. Soient $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. L'expression

$$\text{Cov}[X, Y] := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \quad (1.14)$$

est appelée la *covariance* de X et Y . ◊

Remarque 1. $\text{Cov}[X, Y] = \langle X - \mathbb{E}[X], Y - \mathbb{E}[Y] \rangle$, où $\langle \cdot, \cdot \rangle$ est le produit scalaire sur $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Remarque 2. Si X et Y sont indépendantes, il suit du théorème 1.9 que leur covariance est nulle. Cependant, la réciproque n'est pas vraie.

Remarque 3. L'inégalité de Cauchy–Schwarz sur $L^2(\Omega, \mathcal{F}, \mathbb{P})$ implique que pour deux variables aléatoires $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$,

$$|\text{Cov}[X, Y]|^2 \leq \|X - \mathbb{E}[X]\|_2^2 \|Y - \mathbb{E}[Y]\|_2^2 = \text{Var}[X]\text{Var}[Y]. \quad (1.15)$$

Il arrive qu'il soit plus pratique d'utiliser un analogue adimensionnel de la covariance.

Définition 1.20. Soient $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. L'expression

$$\text{Corr}[X, Y] := \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X]\text{Var}[Y]}} \quad (1.16)$$

est appelée la *corrélation* entre X et Y . \diamond

Remarque. Il suit de la remarque 3 sur la définition 1.19 de la covariance que la corrélation entre deux variables aléatoires est toujours bornée entre -1 et 1 .

De façon plus générale, les relations entre différentes variables aléatoires peuvent être étudiées à partir de leurs moments.

Définition 1.21. Soient $X_1, \dots, X_n \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ et $k_1, \dots, k_n \in \mathbb{N}_0$ tels que $k := k_1 + \dots + k_n \leq p$. On dit que $\mathbb{E}[X_1^{k_1} \cdots X_n^{k_n}]$ est un *moment d'ordre k* ou un *k -ième moment* des variables aléatoires X_1, \dots, X_n . De plus, on dit que $\mathbb{E}[(X_1 - \mathbb{E}[X_1])^{k_1} \cdots (X_n - \mathbb{E}[X_n])^{k_n}]$ est un *moment centré d'ordre k* ou un *k -ième moment centré* des variables aléatoires X_1, \dots, X_n . \diamond

Remarque. L'espérance d'une variable aléatoire est son premier moment et sa variance est son second moment centré. De même, la covariance entre deux variables aléatoires est un moment centré d'ordre 2.

L'espérance conditionnelle permet aussi d'étudier les liens entre deux variables aléatoires.

Définition 1.22. Soient $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ et \mathcal{G} une sous- σ -algèbre de \mathcal{F} . Si $X \geq 0$ p.s., l'*espérance conditionnelle de X sachant \mathcal{G}* est une variable aléatoire $\mathbb{E}[X | \mathcal{G}]$ telle que :

- i) $\mathbb{E}[X | \mathcal{G}]$ est \mathcal{G} -mesurable;
- ii) Pour tout $A \in \mathcal{G}$, $\mathbb{E}[\mathbf{1}_A X] = \mathbb{E}[\mathbf{1}_A \mathbb{E}[X | \mathcal{G}]]$.

Pour le cas général, on écrit $X = X^+ - X^-$ où X^+ et X^- sont positives, et on définit l'espérance conditionnelle de X sachant \mathcal{G} comme $\mathbb{E}[X | \mathcal{G}] := \mathbb{E}[X^+ | \mathcal{G}] - \mathbb{E}[X^- | \mathcal{G}]$. \diamond

Remarque. Il est non trivial l'espérance conditionnelle est même bien définie. Cependant, on peut montrer qu'elle existe et est unique p.s., par exemple à partir du théorème de Radon–Nikodym.

Cette définition peut être étendue à une relation entre deux variables aléatoires.

Définition 1.23. Soient $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. On appelle $\mathbb{E}[X | \sigma(Y)]$ l'*espérance conditionnelle de X sachant Y* , et on note aussi $\mathbb{E}[X | Y] := \mathbb{E}[X | \sigma(Y)]$. \diamond

L'espérance conditionnelle permet de généraliser la notion de probabilité conditionnelle. Avant d'en donner la définition générale, on rappelle d'abord la définition élémentaire de la probabilité conditionnelle d'un événement en sachant un autre.

Définition 1.24. Soient $A, B \in \mathcal{F}$ avec $\mathbb{P}(B) > 0$. La *probabilité conditionnelle de A sachant B* , notée $\mathbb{P}[A | B]$, est définie par la relation $\mathbb{P}[A | B]\mathbb{P}(B) = \mathbb{P}(A \cap B)$. \diamond

Remarque 1. Il n'est pas difficile de montrer que la fonction $A \mapsto \mathbb{P}[A | B]$ est une mesure de probabilité sur (Ω, \mathcal{F}) .

Remarque 2. Il peut être pratique de pouvoir considérer la quantité $\mathbb{P}[A | B]$ lorsqu'on ne connaît pas la probabilité de B . Dans ces cas, on peut voir $\mathbb{P}[A | B]$ comme n'importe quelle valeur dans $[0, 1]$ qui satisfait la relation $\mathbb{P}[A | B]\mathbb{P}(B) = \mathbb{P}(A \cap B)$. En effet, si $\mathbb{P}(B) = 0$, alors on a aussi que $\mathbb{P}(A \cap B) = 0$ donc cette relation reste vraie si l'on donne n'importe quelle valeur dans $[0, 1]$ à $\mathbb{P}[A | B]$.

Définition 1.25. Soit $A \in \mathcal{F}$. Si \mathcal{G} est une sous- σ -algèbre de \mathcal{F} , l'*espérance conditionnelle $\mathbb{P}[A | \mathcal{G}] := \mathbb{E}[\mathbf{1}_A | \mathcal{G}]$* est appelée la *probabilité conditionnelle de A sachant \mathcal{G}* . De la même façon, si $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, l'*espérance conditionnelle $\mathbb{P}[A | X] := \mathbb{E}[\mathbf{1}_A | X]$* est appelée la *probabilité conditionnelle de A sachant X* . \diamond

On conclut cette sous-section par un résultat des plus importants en théorie des probabilités : la loi des grands nombres.

Théorème 1.10 – Loi des grands nombres. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires à valeurs réelles, indépendantes par paires et identiquement distribuées. Si les X_n ont une espérance $a \in \mathbb{R}$ et une variance finie, alors $\frac{1}{N} \sum_{n=1}^N X_n \rightarrow a$ en probabilité et p.s. \blacklozenge

Remarque 1. Le résultat quant à la convergence en probabilité est la loi *faible* des grands nombres, et celui quant à la convergence p.s. est la loi *forte*.

Remarque 2. Il serait possible d'avoir des hypothèses moins strictes sur les X_n , mais cette version de la loi des grands nombres sera suffisante ici.

§ 1.2.3. Espaces produit et noyaux de transition

Lorsque l'on travaille avec plusieurs espaces de probabilité, il est souvent essentiel de pouvoir les combiner ou les relier entre eux. Une première façon de combiner différents espaces en un seul est d'en prendre le produit.

Définition 1.26. Soit T un ensemble. Pour chaque $t \in T$, soient $(\Omega_t, \mathcal{F}_t)$ un espace mesurable et $A_t \in \mathcal{F}_t$. Le produit cartésien $\prod_{t \in T} A_t := \{\omega \in \prod_{t \in T} \Omega_t : \forall t \in T, \omega(t) \in A_t\}$ est appelé un *rectangle mesurable* dans $\prod_{t \in T} \Omega_t$ si A_t est différent de Ω_t seulement pour un nombre fini de $t \in T$. On note $\bigotimes_{t \in T} \mathcal{F}_t$ la σ -algèbre engendrée par les rectangles mesurables dans $\prod_{t \in T} \Omega_t$. L'espace mesurable $(\prod_{t \in T} \Omega_t, \bigotimes_{t \in T} \mathcal{F}_t)$ est appelé le *produit* des $(\Omega_t, \mathcal{F}_t)$. \diamond

Remarque. Lorsque $(\Omega_t, \mathcal{F}_t) = (\Omega, \mathcal{F})$ pour tout $t \in T$, on note aussi $(\Omega, \mathcal{F})^T$ ou $(\Omega^T, \mathcal{F}^T)$ pour le produit des $(\Omega_t, \mathcal{F}_t)$.

Dans le cas où l'ensemble d'indices T est fini, il est assez simple de définir une mesure sur l'espace produit à partir de mesures sur chacun des espaces initiaux.

Théorème 1.11. Pour chaque $j \in \{1, \dots, n\}$, soit un espace de probabilité $(\Omega_j, \mathcal{F}_j, \mathbb{P}_j)$. Alors, il existe une unique mesure de probabilité \mathbb{P} sur $(\prod_{j=1}^n \Omega_j, \bigotimes_{j=1}^n \mathcal{F}_j)$ telle que pour $A_j \in \mathcal{F}_j$, $\mathbb{P}(\prod_{j=1}^n A_j) = \prod_{j=1}^n \mathbb{P}_j(A_j)$. On note $\mathbb{P} = \bigotimes_{j=1}^n \mathbb{P}_j$. \blacklozenge

La question de définir une mesure produit sur une collection infinie d'espaces mesurables est cependant plus épineuse. Des résultats qui permettent de définir de telles mesures seront donnés plus loin, lorsqu'on abordera les processus stochastiques.

Une autre façon importante de relier différents espaces entre eux est à travers des noyaux intégraux, qu'on appelle des noyaux de transition. Ces objets permettent notamment de définir des mesures qui peuvent dépendre d'un paramètre.

Définition 1.27. Soient deux espaces mesurables (E, \mathcal{E}) et (F, \mathcal{F}) . Un *noyau de transition de (E, \mathcal{E}) dans (F, \mathcal{F})* est une fonction $K: E \times \mathcal{F} \rightarrow [0, \infty]$ telle que :

- i) $\forall x \in E, A \mapsto K(x, A)$ est une mesure sur (F, \mathcal{F}) ;
- ii) $\forall A \in \mathcal{F}, x \mapsto K(x, A)$ est \mathcal{E} -mesurable.

Des noyaux de transition peuvent être combinés à partir de leur produit.

Définition 1.28. Soient trois espaces mesurables (E, \mathcal{E}) , (F, \mathcal{F}) et (G, \mathcal{G}) et deux noyaux de transition $K: E \times \mathcal{F} \rightarrow [0, \infty]$ et $L: F \times \mathcal{G} \rightarrow [0, \infty]$. La fonction $KL: E \times \mathcal{G} \rightarrow [0, \infty]$ définie par

$$KL(x, B) := \int_F K(x, dy)L(y, B) \quad (1.17)$$

est appelée le *produit* de K et de L . \diamond

Remarque. Le théorème de la convergence monotone et les propriétés des noyaux transition permettent de montrer qu'un tel produit est aussi un noyau de transition.

On s'intéressera particulièrement à deux cas de noyaux de transition. D'abord, dans le contexte des processus de Markov, le cas particulier suivant sera très important.

Définition 1.29. Soit un espace mesurable (E, \mathcal{E}) . Un *noyau de Markov sur (E, \mathcal{E})* est un noyau de transition K de (E, \mathcal{E}) dans lui-même tel que pour tout $x \in E$, $K(x, E) = 1$. \diamond

Le cas suivant donne le point de départ d'une autre méthode qui permet de construire des mesures de probabilité sur des espaces produit.

Définition 1.30. Soit (E, \mathcal{E}, μ) un espace de probabilité et soit (F, \mathcal{F}) un espace mesurable. Une mesure de probabilité aléatoire ν de (E, \mathcal{E}, μ) dans (F, \mathcal{F}) est un noyau de transition de (E, \mathcal{E}) dans (F, \mathcal{F}) tel que $\nu(x, F) = 1$ pour tout $x \in E$. \diamond

Remarque. De façon équivalente, il est aussi possible de définir une mesure aléatoire comme une variable aléatoire à valeurs dans un espace de fonctions, ce qui est fait par exemple dans [62]. On ne donnera pas davantage de détails sur ce point de vue, mais il est intéressant de remarquer que de cette approche rend la terminologie plus claire : une mesure de probabilité aléatoire est bien une variable aléatoire à proprement parler.

Le théorème suivant donne la méthode générale pour obtenir une mesure de probabilité sur un espace produit.

Théorème 1.12. Soit (E, \mathcal{E}, μ) un espace de probabilité et (F, \mathcal{F}) un espace mesurable. Soit ν une mesure de probabilité aléatoire de (E, \mathcal{E}, μ) dans (F, \mathcal{F}) . Alors il existe une mesure de probabilité \mathbb{P} sur l'espace produit $(E \times F, \mathcal{E} \otimes \mathcal{F})$ telle que si X est une variable aléatoire sur $E \times F$, alors

$$\mathbb{E}[X] = \int_{E \times F} X \, d\mathbb{P} = \int_E \int_F X(x, y) \nu(x, dy) \mu(dx). \quad (1.18)$$

◆

Remarque 1. La relation (1.18) décrit aussi la mesure \mathbb{P} , car $\mathbb{P}(A) = \mathbb{E}[\mathbf{1}_A]$ pour $A \in \mathcal{E} \otimes \mathcal{F}$.

Remarque 2. Le cas où $n = 2$ au théorème 1.11 est un cas particulier du théorème 1.12, en prenant $(E, \mathcal{E}, \mu) = (\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$, $(F, \mathcal{F}) = (\Omega_2, \mathcal{F}_2)$ et la mesure aléatoire $\nu(x, A) := \mathbb{P}_2(A)$.

Dans le contexte du théorème précédent, une fonction mesurable Y définie sur (F, \mathcal{F}) peut naturellement être vue comme une variable aléatoire $(x, y) \mapsto Y(y)$ sur l'espace de probabilité produit. Pour chaque $x \in E$, il est alors possible de définir l'espérance de Y à partir de la mesure de probabilité $A \mapsto \nu(x, A)$ sur (F, \mathcal{F}) . La proposition suivante donne une interprétation de ces espérances.

Proposition 1.13. Soit ν une mesure de probabilité aléatoire d'un espace de probabilité (E, \mathcal{E}, μ) dans un espace mesurable (F, \mathcal{F}) . Soit \mathbb{P} la mesure de probabilité sur $(E \times F, \mathcal{E} \otimes \mathcal{F})$ introduite au théorème 1.12. Si $Y: F \rightarrow \mathbb{R}$ est \mathcal{F} -mesurable, alors pour tout $(x, y) \in E \times F$,

$$\mathbb{E}[Y | \pi_E](x, y) = \int_F Y(z) \nu(x, dz), \quad (1.19)$$

où $\pi_E: E \times F \rightarrow E$ note la projection sur E . \diamond

Remarque 1. On utilise ici un léger abus de notation : on voit Y comme une fonction sur F mais aussi comme la fonction $(x, y) \mapsto Y(y)$ sur $E \times F$.

Remarque 2. La σ -algèbre $\sigma(\pi_E)$ contient toute l'information disponible sur l'espace (E, \mathcal{E}) , mais aucune information sur (F, \mathcal{F}) . L'espérance conditionnelle $\mathbb{E}[Y | \pi_E](x, y)$, qui ne dé-

pend pas de y , correspond alors simplement à l'espérance de Y sur l'espace de probabilité (F, \mathcal{F}, ν^x) , où $\nu^x(A) := \nu(x, A)$. Étant donnée cette interprétation, on notera généralement $\mathbb{E}[Y | x]$ plutôt que $\mathbb{E}[Y | \pi_E](x, y)$.

Démonstration. On remarque d'abord que si $Y = \sum_{j=1}^n c_j \mathbb{1}_{A_j}$ pour des $c_j > 0$ et des $A_j \in \mathcal{F}$, alors $x \mapsto \int_F Y(z) \nu(x, dz) = \sum_{j=1}^n c_j \nu(x, A_j)$ est \mathcal{E} -mesurable, puisque chacune des fonctions $x \mapsto \nu(x, A_j)$ est \mathcal{E} -mesurable. La fonction $x \mapsto \int_F Y(z) \nu(x, dz)$ est donc toujours \mathcal{E} -mesurable par le théorème de la convergence monotone.

On note alors que $\sigma(\pi_E) = \{A \times F : A \in \mathcal{E}\}$. La fonction $\bar{Y}(x, y) := \int_F Y(z) \nu(x, dz)$ est donc $\sigma(\pi_E)$ -mesurable, car elle ne dépend pas de la seconde coordonnée. Si $A \times F \in \sigma(\pi_E)$, alors

$$\begin{aligned} \mathbb{E}[\mathbb{1}_{A \times F} Y] &= \int_E \int_F \mathbb{1}_{A \times F}(x, y) Y(y) \nu(x, dy) \mu(dx) = \int_E \mathbb{1}_A(x) \int_F Y(y) \nu(x, dy) \int_F \nu(x, dz) \mu(dx) \\ &= \int_E \int_F \mathbb{1}_{A \times F}(x, z) \bar{Y}(x, z) \nu(x, dz) \mu(dx) = \mathbb{E}[\mathbb{1}_{A \times F} \bar{Y}]. \end{aligned} \quad (1.20)$$

Ainsi, on voit que $\bar{Y} = \mathbb{E}[Y | \pi_E]$ \mathbb{P} -p.s. ■

§ 1.2.4. Processus stochastiques

Un processus stochastique est simplement la donnée d'une collection de variables aléatoires.

Définition 1.31. Soient T un ensemble, $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et (E, \mathcal{E}) un espace mesurable. Une collection $\{X_t\}_{t \in T}$ de variables aléatoires $X_t: \Omega \rightarrow E$ est appelée un *processus stochastique* sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans (E, \mathcal{E}) . ◇

L'ensemble T de la définition précédente est généralement interprété comme un ensemble de temps, qui peut être continu (par exemple $T = [0, \infty)$) ou discret (par exemple $T = \mathbb{N}$). Un processus stochastique est ainsi souvent utilisé pour représenter l'état d'un système qui évolue de façon aléatoire.

Exemple 1.4. Soient (E, \mathcal{E}) un espace mesurable et $(\Omega, \mathcal{F}) := (E, \mathcal{E})^T$. Pour chaque $t \in T$, soit $X_t: \Omega \rightarrow E$ définie par $X_t(\omega) := \omega(t)$. Pour $A \in \mathcal{E}$, $X_t^{-1}(A) = \{\omega \in \Omega : \omega(t) \in A\}$, qui est un rectangle mesurable sur $(E, \mathcal{E})^T$. Chaque X_t est donc bien une variable aléatoire. On appelle le processus stochastique $\{X_t\}_{t \in T}$ ainsi défini un *processus de projection de coordonnées*. En fait, puisque chaque rectangle mesurable sur $(E, \mathcal{E})^T$ est une intersection finie de préimages $X_t^{-1}(A_t)$ pour certains $A_t \in \mathcal{E}$, on voit même que $\mathcal{E}^T = \sigma(X_t : t \geq 0)$. ▲

Le processus de projection de coordonnées introduit à l'exemple précédent peut à première vue sembler être un cas très particulier, mais il suffit en fait à définir de nombreux processus stochastiques différents. En effet, la distinction entre différents processus réside autant, voire davantage, dans les espaces de probabilité sur lesquels ils sont définis, et en particulier dans les mesures de probabilité, plutôt que sur les variables aléatoires qui composent le processus. Ainsi, lorsqu'on parle de construire un processus stochastique, on parle sou-

vent de construire une mesure de probabilité sur un certain espace mesurable de façon à ce qu'une certaine collection de fonctions mesurables déjà connues devienne un processus stochastique avec certaines propriétés, plutôt que de construire de nouvelles variables aléatoires sur un espace de probabilité déjà connu.

Un théorème important, le théorème d'extension de Kolmogorov, garantit sous des conditions assez faibles l'existence de mesures de probabilité qui permettent de définir plusieurs processus stochastiques. Pour introduire ce théorème, on introduit préalablement quelques nouveaux objets.

Définition 1.32. Un espace mesurable (E, \mathcal{E}) est dit *standard* s'il existe une bijection mesurable d'inverse mesurable entre (E, \mathcal{E}) et $(B, \mathcal{B}(B))$ pour un certain $B \in \mathcal{B}(\mathbb{R})$. \diamond

Remarque. Lorsqu'il existe une telle bijection entre (E, \mathcal{E}) et $(B, \mathcal{B}(B))$, on dit que ces deux espaces mesurables sont *isomorphes*.

Exemple 1.5. Si E est fini, alors $(E, 2^E)$ est standard. En effet, si n est le nombre d'éléments que comporte E , une bijection entre E et $\{1, \dots, n\}$ peut être définie simplement en énumérant les éléments de E . Comme ces deux espaces sont discrets, cette bijection est nécessairement mesurable et d'inverse mesurable. \blacktriangleleft

Définition 1.33. Soient T un ensemble et (E, \mathcal{E}) un espace mesurable standard. Pour tout $J \subset T$ fini, soit π_J une mesure de probabilité sur $(E, \mathcal{E})^J$. Si $K \subset J$, soit $\text{proj}_{JK}: E^J \rightarrow E^K$ la projection naturelle. On dit que $\{\pi_J : J \subset T \text{ est fini}\}$ est une *famille cohérente de mesures de probabilité* sur $(E, \mathcal{E})^T$ si pour tout $J \subset T$ fini et pour tout $K \subset J$, $\pi_K = \pi_J \circ \text{proj}_{JK}^{-1}$. \diamond

On peut maintenant énoncer le théorème, un résultat classique très important.

Théorème 1.14 – Théorème d'extension de Kolmogorov. Soient T un ensemble et (E, \mathcal{E}) un espace mesurable standard. Soit $\{\pi_J : J \subset T \text{ est fini}\}$ une famille cohérente de mesures de probabilité sur $(E, \mathcal{E})^T$. Pour chaque $t \in T$, soit $X_t: E^T \rightarrow E$ définie par $X_t(\omega) := \omega(t)$. Alors, il existe une mesure de probabilité \mathbb{P} sur $(E, \mathcal{E})^T$ telle que pour tout $J \subset T$ fini, $\mathbb{P}[X_J \in A] = \pi_J(A)$, où $X_J := (X_t)_{t \in J}$. \blacklozenge

On s'intéressera particulièrement aux processus de Markov en temps continu. L'idée de ce type de processus est que, connaissant l'état X_s du processus en un temps s , la connaissance des états antérieurs à s n'apporte pas davantage d'information sur les états futurs X_t pour $t > s$. Pour bien définir ce type de processus, on introduit d'abord la notion de filtration. Pour simplifier la discussion, on prendra $[0, \infty)$ comme ensemble de temps.

Définition 1.34. Soit $\{X_t\}_{t \geq 0}$ un processus stochastique sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Pour chaque $t \geq 0$, soit $\mathcal{F}_t^X := \sigma(X_s : s < t)$. La collection $\{\mathcal{F}_t^X\}_{t \geq 0}$ est appelée la *filtration engendrée par $\{X_t\}_{t \geq 0}$* . \diamond

La notion de processus de Markov peut alors être introduite.

Définition 1.35. Soit $\{X_t\}_{t \geq 0}$ un processus stochastique sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans un espace (E, \mathcal{E}) . Si $\mathbb{P}[X_t \in A | \mathcal{F}_s^X] = \mathbb{P}[X_t \in A | X_s]$ pour tout $A \in \mathcal{E}$ et pour $s, t \geq 0$ avec $t > s$, alors on dit que $\{X_t\}_{t \geq 0}$ est un *processus de Markov*. ◇

Pour construire un espace de probabilité sur lequel définir un processus de Markov, on peut utiliser le théorème d'extension de Kolmogorov. Pour ce faire, le plus simple est d'introduire un semi-groupe de transition.

Définition 1.36. Soit (E, \mathcal{E}) un espace mesurable. Un *semi-groupe (de transition) de Markov* sur (E, \mathcal{E}) est une collection $\{K_{s,t} : t > s \geq 0\}$ de noyaux de Markov sur (E, \mathcal{E}) telle que pour $s, t, u \geq 0$ avec $s < t < u$, $K_{s,t}K_{t,u} = K_{s,u}$. ◇

Remarque 1. On interprète E comme un ensemble d'états possibles pour un système. Pour un $x \in E$ et un $A \in \mathcal{E}$, on interprète alors $K_{s,t}(x, A)$ comme représentant la probabilité que l'état du système se trouve dans A au temps t sachant qu'il est dans l'état x au temps s .

Remarque 2. L'équation $K_{s,t}K_{t,u} = K_{s,u}$ est appelée *l'équation de Chapman–Kolmogorov*.

Étant donné un semi-groupe de Markov sur un espace mesurable (E, \mathcal{E}) , il est possible de construire une mesure de probabilité sur l'espace produit $(E, \mathcal{E})^{[0,\infty)}$.

Théorème 1.15. Soient (E, \mathcal{E}) un espace mesurable standard et $\{K_{s,t} : t > s \geq 0\}$ un semi-groupe de Markov sur (E, \mathcal{E}) . Pour chaque $t \geq 0$, soit $X_t : E^{[0,\infty)} \rightarrow E$ définie par $X_t(\omega) := \omega(t)$. Soit μ une mesure de probabilité sur (E, \mathcal{E}) . Alors il existe une mesure de probabilité \mathbb{P} sur $(E, \mathcal{E})^{[0,\infty)}$ telle que si $A \in \mathcal{E}$ et $t > s \geq 0$, $\mathbb{P}[X_0 \in A] = \mu(A)$ et $\mathbb{P}[X_t \in A | \mathcal{F}_s^X] = \mathbb{P}[X_t \in A | X_s] = K_{s,t}(X_s, A)$. ◆

L'idée de la démonstration est de prendre, pour $J = (t_1, \dots, t_n) \subset [0, \infty)$,

$$\pi_J(A) := \int_A \left(\prod_{j=1}^n K_{t_{j-1}, t_j}(x_{j-1}, dx_j) \right) \mu(dx_0). \quad (1.21)$$

Les π_J forment alors une famille cohérente de mesure de probabilité sur $(E, \mathcal{E})^{[0,\infty)}$. Ainsi, le théorème d'extension de Kolmogorov garantit l'existence d'une mesure \mathbb{P} sur $(E, \mathcal{E})^{[0,\infty)}$ avec les propriétés voulues. La propriété de Markov $\mathbb{P}[X_t \in A | \mathcal{F}_s^X] = \mathbb{P}[X_t \in A | X_s]$ suit ensuite des propriétés de l'espérance conditionnelle. Les détails sont donnés dans [18].

§ 1.2.5. Chaînes de Markov à temps continu homogène

Dans la suite, les processus auxquels on s'intéressera particulièrement sont des chaînes de Markov à temps continu homogène, c'est-à-dire des processus de Markov à temps continu dont l'espace des états est discret et dont les noyaux de transition $K_{s,t}$ ne dépendent que de la différence $t - s$. Cette sous-section présente certains résultats généraux qui permettent de construire et d'étudier de tels processus. Les résultats donnés ici sont basés sur les livres de Doob [25] et de Norris [74].

Soient E un ensemble fini et $\mathcal{E} := 2^E$. À nouveau, on interprète E comme l'ensemble des

états possibles d'un système. Une chaîne de Markov qui représente l'évolution de ce système peut être construite à partir d'un générateur.

Définition 1.37. Soient $q := E \times E \rightarrow \mathbb{R}$ et $Q := \{q(x, y)\}_{x, y \in E}$. On dit que Q est le *générateur* d'une chaîne de Markov en temps continu homogène si :

- i) $\forall x, y \in E$ avec $x \neq y$, $q(x, y) \geq 0$;
- ii) $\forall x \in E$, $q(x, x) \leq 0$;
- iii) $\forall x \in E$, $\sum_{y \in E} q(x, y) = 0$.

◇

Remarque 1. On interprète $q(x, y)$ comme le taux de transition de l'état x vers l'état y .

Remarque 2. Puisque E est fini, on voit Q comme une matrice dont les éléments sont les $q(x, y)$.

Un générateur $Q := \{q(x, y)\}_{x, y \in E}$ suffit à construire un processus de Markov. En effet, pour $t \geq 0$, soient $p_t(x, y)$ les éléments de la matrice $P_t := e^{tQ}$. À partir de manipulations matricielles et de propriétés de l'exponentielle, on peut montrer que les P_t possèdent les propriétés suivantes :

- i) $\forall t \geq 0$, $p_t(x, y) \geq 0$;
- ii) $\forall t \geq 0$, $\forall x \in E$, $\sum_{y \in E} p_t(x, y) = 1$;
- iii) $\forall s, t \geq 0$, $P_t P_s = P_{t+s}$, c'est-à-dire que $\forall x, y \in E$,

$$\sum_{z \in E} p_t(x, z) p_s(z, y) = p_{t+s}(x, y). \quad (1.22)$$

Pour $s, t \geq 0$, on définit alors $K_{s,t}: E \times \mathcal{E} \rightarrow [0, 1]$ par $K_{s,t}(x, A) := \sum_{y \in A} p_{t-s}(x, y)$. Il est facile de vérifier que les $K_{s,t}$ ainsi définis forment un semi-groupe de Markov. En effet, puisque (E, \mathcal{E}) est discret, toutes les fonctions $x \mapsto K_{s,t}(x, A)$ sont forcément mesurables. De plus, pour chaque $x \in E$, la fonction $A \mapsto K_{s,t}(x, A)$ est additive par définition et vaut zéro en \emptyset , donc il s'agit d'une mesure sur (E, \mathcal{E}) . La propriété ii) des P_t équivaut ensuite à dire que chaque mesure $A \mapsto K_{s,t}(x, A)$ est une mesure de probabilité. Chaque $K_{s,t}$ est donc bien un noyau de Markov sur (E, \mathcal{E}) .

Pour que les $K_{s,t}$ forment un semi-groupe de Markov, il reste à montrer qu'ils respectent l'équation de Chapman–Kolmogorov. Soient $x \in E$ et $A \in \mathcal{E}$. Pour $s, t, u \geq 0$,

$$K_{s,t} K_{t,u}(x, A) = \int_E K_{s,t}(x, dy) K_{t,u}(y, A) = \sum_{y \in E} K_{s,t}(x, \{y\}) K_{t,u}(y, A) \quad (1.23)$$

$$= \sum_{y \in E} p_{t-s}(x, y) \sum_{z \in A} p_{u-t}(y, z) = \sum_{z \in A} \sum_{y \in E} p_{t-s}(x, y) p_{u-t}(y, z). \quad (1.24)$$

La propriété iii) donne alors directement que

$$K_{s,t} K_{t,u}(x, A) = \sum_{z \in A} p_{u-s}(z, z) = K_{s,u}(z, A). \quad (1.25)$$

Ainsi, les $K_{s,t}$ respectent bien l'équation de Chapman–Kolmogorov $K_{s,t} K_{t,u} = K_{s,u}$. Ils forment donc un semi-groupe de Markov.

L'espace (E, \mathcal{E}) est fini et discret. Il est donc standard d'après l'exemple 1.5. Le théorème 1.15 montre alors que pour toute mesure de probabilité μ sur (E, \mathcal{E}) , il existe une mesure de probabilité \mathbb{P} sur $(E, \mathcal{E})^{[0, \infty)}$ telle que si $\{X_t\}_{t \geq 0}$ est le processus de projection de coordonnées sur $(E, \mathcal{E})^{[0, \infty)}$, défini par $X_t(\omega) := \omega(t)$, alors pour tout $A \in \mathcal{E}$ et pour $t > s \geq 0$, $\mathbb{P}[X_0 \in A] = \mu(A)$ et

$$\mathbb{P}[X_t \in A | \mathcal{F}_s^X] = \mathbb{P}[X_t \in A | X_s] = K_{s,t}(X_s, A) = \sum_{y \in A} p_{t-s}(X_s, y). \quad (1.26)$$

Cette propriété de la mesure \mathbb{P} permet d'obtenir de nombreuses relations très utiles. Par exemple, puisque E est discret, on a pour $A \in E$ que

$$\mathbb{P}[X_t \in A] = \sum_{x \in E} \mathbb{P}[X_t \in A | X_0 = x] \mathbb{P}[X_0 = x] = \sum_{x \in E} K_{0,t}(x, A) \mu(\{x\}), \quad (1.27)$$

donc en explicitant l'expression de $K_{0,t}$, on voit que

$$\mathbb{P}[X_t \in A] = \sum_{x \in E} \sum_{y \in A} \mu(\{x\}) p_t(x, y). \quad (1.28)$$

De plus, d'après la définition des p_t et d'après les propriétés de l'exponentielle, on voit également que lorsque $\Delta t \downarrow 0$, pour tout $x, y \in E$,

$$\mathbb{P}[X_{t+\Delta t} = y | X_t = x] = K_{t,t+\Delta t}(x, \{y\}) = p_{\Delta t}(x, y) = \delta_{xy} + q(x, y)\Delta t + o(\Delta t), \quad (1.29)$$

où δ_{xy} est le symbole de Kronecker, qui vaut 1 si $x = y$ et 0 sinon.

Finalement, on voit que la matrice $P_t = e^{tQ}$ est dérivable selon t , puisque la série (1.4) qui définit cette fonction converge uniformément sur n'importe quel intervalle borné. Un calcul direct donne que $\dot{P}_t = Qe^{tQ} = QP_t$, où le point note la dérivée par rapport à t . Le résultat

$$\dot{P}_t = QP_t \quad (1.30)$$

s'appelle l'*équation de Kolmogorov*. Ce résultat montre que pour tout $A \in \mathcal{E}$, la probabilité $\mathbb{P}[X_t \in A]$ est dérivable selon t . En effet, on a vu que

$$\mathbb{P}[X_t \in A] = \sum_{x \in E} \sum_{y \in A} \mu(\{x\}) p_t(x, y). \quad (1.31)$$

Ainsi, puisque P_t est dérivable selon t , chacune de ses composantes $p_t(x, y)$ l'est aussi, donc la somme $\mathbb{P}[X_t \in A]$ l'est aussi.

§ 1.3. Problèmes de fermeture de moments

L'étude d'un processus stochastique dont la loi de probabilité n'est que partiellement connue est difficile. Parfois, il arrive que l'on ait une description de l'évolution d'un système sous la forme d'un système d'équations différentielles qui décrit les moments d'un processus

stochastique. Dans ce cas, peut-on obtenir une description, possiblement approximative, de l'évolution du système en termes de seulement quelques-uns de ces moments ? Il est possible de le faire en trouvant une solution à un problème de fermeture de moments. On présente ici un survol de ce type de problème. Pour une introduction plus complète, on réfère à l'excellente revue de Kuehn [65].

Pour énoncer précisément le problème, on considère $\{X_t\}_{t \geq 0}$ un processus stochastique à valeurs dans \mathbb{R}^d et défini sur un espace de probabilité quelconque. On énumère les moments des composantes de X_t à partir d'un ensemble d'indices \mathcal{J} , et on les note par des fonctions $m_j: [0, \infty) \rightarrow \mathbb{R}$ pour $j \in \mathcal{J}$. Puisque ces moments sont en quantité dénombrable, on identifie $\mathcal{J} = \mathbb{N}$ afin de simplifier la notation de la suite.

On suppose ensuite que l'évolution des m_j est régie par un système d'équations différentielles de la forme

$$\dot{m}_j = f_j(m_1, m_2, m_3, m_4, \dots) \quad (1.32)$$

pour $j \in \mathbb{N}$ et $f_j: \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$, où le point note la dérivée. Puisqu'il existe une infinité de moments, le système (1.32) a en fait une infinité de dimensions. La tâche de réduire un tel système à un système autonome de dimension finie s'appelle un *problème de fermeture de moments*¹. Trouver une solution à ce problème consiste à choisir un certain $k \in \mathbb{N}$ et à trouver une certaine fonction $F: \mathbb{R}^k \rightarrow \mathbb{R}^{\mathbb{N}}$ telle que

$$F(m_1, \dots, m_k) = (m_{k+1}, m_{k+2}, m_{k+3}, \dots). \quad (1.33)$$

L'évolution des k premiers moments est alors donnée par le système

$$\dot{m}_j = f_j(m_1, \dots, m_k, F(m_1, \dots, m_k)), \quad (1.34)$$

qui est autonome. On dit alors que le système est *fermé*.

Les problèmes de fermeture de moments sont peu étudiés dans un cadre général, mais des cas précis surviennent et sont étudiés dans de nombreuses disciplines, notamment en biologie mathématique et en physique statistique [65]. Par exemple, comme on le verra en détail au chapitre 3, des fermetures de moments sont utilisées dans la construction de modèles compartimentaux [60] en épidémiologie.

Le premier défi d'un problème de fermeture de moments comme celui décrit plus haut est de trouver la fonction F donnée en (1.33) qui exprime les moments d'ordre supérieur en fonction de certains moments donnés. Typiquement, les idées utilisées pour trouver une telle fonction sont intuitives ou basées sur des suppositions supplémentaires [65]. Ces arguments dépendent souvent de la forme du processus stochastique ou bien du phénomène qu'il représente.

1. Traduction libre de « *moment closure* », l'expression généralement utilisée dans la littérature anglophone.

Une fermeture simple, qui peut toujours être utilisée, consiste à négliger les moments centrés à partir d'un certain ordre ℓ . Ceci correspond simplement à choisir k comme le dernier moment d'ordre $\ell - 1$, puis à prendre $F \equiv 0$. On utilise souvent le qualificatif « *de champ moyen* » pour référer au cas $\ell = 2$ de cette fermeture, où seuls les moments d'ordre 1 sont conservés ; on dit dans ce cas que le système fermé (1.34) est un modèle de champ moyen [60, 82].

Il est très fréquent que la fermeture choisie ne soit valide que si certaines suppositions additionnelles sont respectées. La difficulté majeure d'un problème de fermeture de moments est ainsi de déterminer dans quelle mesure les solutions du système fermé (1.34) sont une bonne approximation des véritables moments du processus stochastique. Cette question est très difficile à étudier rigoureusement [65, 93]. En pratique, pour déterminer si une fermeture de moments donne de bons résultats, on doit souvent se contenter de comparer des trajectoires du processus stochastique initial — qui peuvent être obtenues à partir de simulations numériques — aux solutions du système fermé (1.34).

Intuitivement, on s'attend à ce que le système fermé fournisse une description plus précise lorsque davantage de moments sont considérés. Ceci a déjà été vérifié pour plusieurs modèles. Par exemple, pour des modèles épidémiologiques, certains auteurs dont [71, 54] ont comparé des simulations numériques de processus stochastiques à des solutions de systèmes fermés issus de différentes fermetures, et ont conclu que la précision des modèles augmentait avec l'ordre des moments inclus. Cependant, d'autres auteurs émettent des réserves, puisque les propriétés des systèmes dynamiques peuvent changer lorsque davantage de moments sont considérés. Par exemple, il peut arriver que des fermetures qui semblent être basées sur des suppositions sensées mènent à des solutions dont les valeurs n'ont pas de sens par rapport aux moments qu'elles devraient représenter [16]. On verra d'ailleurs un exemple d'un tel phénomène à la section 4.4.2.4.

La recherche d'une bonne solution à un problème de fermeture de moments est ainsi toujours un problème difficile, et dépend de la forme du système (1.32). On étudiera davantage ces idées pour des systèmes donnés aux chapitres 3 et 4.

Chapitre 2

Dynamique markovienne ternaire cyclique sur graphes aléatoires

On aborde dans ce chapitre l'étude de processus stochastiques qui caractérisent les états des nœuds d'un graphe et qui décrivent leur évolution. On étudie également le passage d'un tel modèle microscopique vers un modèle qui décrit la dynamique sur le graphe d'un point de vue macroscopique par un système d'équations différentielles. L'objectif est de donner tous les détails de la construction mathématique, de façon à bien en comprendre les fondements et à mieux cerner les difficultés qui surviennent lors du passage du microscopique au macroscopique.

On se restreint à des processus où les états des noeuds peuvent prendre trois valeurs, qui évoluent de façon cyclique. Cette contrainte est choisie afin de pouvoir modéliser le comportement d'un réseau de neurones biologiques. Cette application sera étudiée en détail au chapitre 4. Les constructions présentées dans ce chapitre seront tout de même assez générales pour pouvoir modéliser d'autres phénomènes. En particulier, on verra au chapitre 3 qu'elles peuvent être appliquées afin de retrouver un modèle classique en épidémiologie.

Dans la première section, on commence par introduire la notion de graphe aléatoire, et on définit certaines structures sur ces graphes. L'idée de travailler avec des graphes aléatoires est un détail technique : on voudra pondérer les liens entre les nœuds, mais on voudra que ces poids soient aléatoires. Ensuite, dans la seconde section, on construit un processus stochastique afin de représenter l'évolution des nœuds d'un graphe aléatoire. Dans la troisième section, on utilise ce processus stochastique comme point de départ afin d'obtenir un système d'équations différentielles qui décrit l'évolution de la probabilité que les nœuds du graphe soient dans un état donné. Afin d'obtenir une description macroscopique du graphe, on divise celui-ci en populations dans la quatrième section, et on y obtient une équation différentielle qui décrit sa dynamique. On verra cependant que cette équation différentielle n'est pas autonome. Dans la dernière section, on verra qu'on peut voir la tâche de la rendre autonome comme un problème de fermeture de moments, et on décrira quelques pistes de solutions à ce problème.

Les modélisations étudiées dans les chapitres suivants dépendent en bonne partie de résultats de ce chapitre. En particulier, on utilise toujours la même notation, introduite à la définition 2.9, pour les états des nœuds d'un graphe. De plus, pour représenter leur évolution, on utilise toujours la même forme de processus stochastique, dont les propriétés sont décrites par le théorème 2.3 et par ses corollaires. Finalement, pour décrire une dynamique sur un graphe d'un point de vue macroscopique, on utilise toujours des équations différentielles de la forme (2.60), qu'on utilise comme point de départ afin d'obtenir un système autonome selon la méthode décrite à la section 2.5.

Les démarches présentées dans ce chapitre sont celles de l'auteur de ce mémoire, qui n'a jamais trouvé de démarches équivalentes dans la littérature. La méthode qu'on utilise à la section 2.3 afin d'obtenir les équations différentielles du point de vue microscopique est inspirée des méthodes classiques utilisées en épidémiologie pour construire des modèles semblables dans le cas où chaque nœud a deux états possibles, qu'on peut trouver notamment dans [94, 95]. Les démarches proposées ici peuvent ainsi être vues comme une généralisation de ces modèles, qui s'applique à davantage de cas.

§ 2.1. Graphes aléatoires

§ 2.1.1. Graphes aléatoires pondérés

Un graphe aléatoire peut être vu simplement comme une variable aléatoire dont les valeurs sont des graphes.

Définition 2.1. Soient $N \in \mathbb{N}$ et $\mathcal{G}_N := \{(V, A) : V = \{1, \dots, N\}, A \subset V \times V\}$ l'ensemble des graphes dirigés à N nœuds. Un *graphe aléatoire d'ordre N* sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$ est une variable aléatoire sur $(\Theta, \mathcal{G}, \mu)$ à valeurs dans l'espace mesurable $(\mathcal{G}_N, \mathcal{2}^{\mathcal{G}_N})$. ◇

Remarque. L'approche usuelle pour définir un graphe aléatoire, adoptée par exemple dans le livre de Bollobás [5], est plutôt de prendre une application identité sur un espace de probabilité construit sur un ensemble de graphes. La mesure de probabilité choisie détermine alors les propriétés du graphe aléatoire. On choisit de ne pas utiliser cette approche afin de garder général l'espace de probabilité utilisé. Ceci permettra de définir sur ce même espace d'autres variables aléatoires qui représenteront diverses propriétés des nœuds.

On s'intéressera uniquement à des graphes aléatoires *pondérés*, c'est-à-dire des graphes aléatoires dans lesquels on pondère l'importance relative des liens. Pour définir un graphe pondéré, il suffit de considérer une matrice carrée dont les entrées caractérisent les liens entre les nœuds, à la manière de la matrice d'adjacence d'un graphe.

Définition 2.2. Soient $N \in \mathbb{N}$ et $w = (w_{jk})_{j,k=1}^N \in \mathbb{R}^{N \times N}$. Soient $V := \{1, \dots, N\}$, et $A \subset V \times V$ l'ensemble des doublets (k, j) tels que $w_{jk} \neq 0$. Dans ce cas, on dit que (V, A) est le graphe induit par w , on appelle w une *matrice de poids*, et on dit que le graphe (V, A) est *pondéré* par w . Un élément w_{jk} de w est appelé le *poids* du lien de k vers j . \diamond

Comme on peut définir un graphe pondéré à partir d'une matrice, on peut définir un graphe aléatoire pondéré à partir d'une matrice aléatoire.

Définition 2.3. Soient $N \in \mathbb{N}$ et $(\Theta, \mathcal{G}, \mu)$ un espace de probabilité. Une *matrice aléatoire de poids* de dimension N sur $(\Theta, \mathcal{G}, \mu)$ est une variable aléatoire $W: \Theta \rightarrow \mathbb{R}^{N \times N}$. \diamond

Remarque. Une matrice aléatoire de poids W de dimension N peut être identifiée à ses composantes $W_{jk}: \Theta \rightarrow \mathbb{R}$, telles que $W(\vartheta) = (W_{jk}(\vartheta))_{j,k=1}^N$. La variable aléatoire W_{jk} est alors interprétée comme le poids du lien du nœud k vers le nœud j .

Exemple 2.1. Si μ_W est une mesure de probabilité sur $(\mathbb{R}^{N \times N}, \mathcal{B}(\mathbb{R}^{N \times N}))$, alors une application identité sur $\mathbb{R}^{N \times N}$ est une matrice aléatoire de poids sur $(\mathbb{R}^{N \times N}, \mathcal{B}(\mathbb{R}^{N \times N}), \mu_W)$. \blacktriangleleft

Pour pouvoir définir un graphe aléatoire pondéré à partir d'une matrice aléatoire, il est nécessaire que l'application qui associe une matrice dans $\mathbb{R}^{N \times N}$ au graphe qu'elle induit soit mesurable. Il n'est pas difficile de montrer que c'est toujours le cas.

Proposition 2.1. Soit $G: \mathbb{R}^{N \times N} \rightarrow \mathcal{G}_N$ l'application telle que pour $w \in \mathbb{R}^{N \times N}$, $G(w)$ est le graphe induit par w . Alors G est mesurable de $(\mathbb{R}^{N \times N}, \mathcal{B}(\mathbb{R}^{N \times N}))$ dans $(\mathcal{G}_N, 2^{\mathcal{G}_N})$. \blacklozenge

Démonstration. Soit un graphe dirigé $g = (V, A) \in \mathcal{G}_N$. Alors $G^{-1}(\{g\}) = \prod_{j,k=1}^N U_{jk}$, où $U_{jk} = \mathbb{R} \setminus \{0\}$ si $(k, j) \in A$ et $U_{jk} = \{0\}$ sinon. Dans ce cas, $U_{jk} \in \mathcal{B}(\mathbb{R})$ pour chaque paire (j, k) , de sorte que $G^{-1}(\{g\}) \in \mathcal{B}(\mathbb{R}^{N \times N})$. Ensuite, si $B \in 2^{\mathcal{G}_N}$, alors $G^{-1}(B) = G^{-1}(\bigcup_{g \in B} \{g\}) = \bigcup_{g \in B} G^{-1}(\{g\})$. Mais puisque \mathcal{G}_N est fini, B est fini. Ainsi, on voit que $G^{-1}(B) = \bigcup_{g \in B} G^{-1}(\{g\}) \in \mathcal{B}(\mathbb{R}^{N \times N})$, et G est mesurable. \blacksquare

Le fait que cette application soit mesurable motive la définition suivante.

Définition 2.4. Soit W une matrice aléatoire de poids de dimension N définie sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$. Soit G l'application introduite à la proposition 2.1. Alors $G_W := G \circ W$ est appelée le *graphe aléatoire pondéré induit par W* . Dans ce cas, on dit aussi que W *pondère* le graphe aléatoire G_W . \diamond

Remarque. Puisque G est mesurable, la composition $G \circ W$ est une variable aléatoire, donc G_W est bien un graphe aléatoire au sens de la définition 2.1.

§ 2.1.2. Paramètres et populations

Afin de pouvoir représenter différentes caractéristiques des nœuds d'un graphe aléatoire, on introduit une notion de paramètre. On considérera toujours les paramètres comme aléatoires.

Définition 2.5. Soit G un graphe aléatoire sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$. Un *paramètre d'ordre k* sur G est une variable aléatoire $\varphi: \Theta \rightarrow \mathbb{R}^{N^k}$. \diamond

Remarque. On interprète un paramètre d'ordre k comme une façon d'associer une valeur réelle à chaque choix de k nœuds du graphe. Ainsi, on identifiera souvent un paramètre à ses composantes : si φ est un paramètre d'ordre k sur un graphe aléatoire d'ordre N , alors on identifie φ à l'ensemble des N^k variables aléatoires $\varphi_{j_1, \dots, j_k}: \Theta \rightarrow \mathbb{R}$ telles que $\varphi(\vartheta) = (\varphi_{j_1, \dots, j_k}(y))_{j_1, \dots, j_k=1}^N$. En particulier, si φ est un paramètre d'ordre 1, on identifie $\varphi = (\varphi_j)_{j=1}^N$. On interprète alors φ_j comme la valeur du paramètre φ pour le nœud j .

Exemple 2.2. Si G est un graphe aléatoire induit par une matrice de poids W , celle-ci est un paramètre d'ordre 2 sur G . \blacktriangleleft

On voudra ensuite pouvoir considérer des graphes dans lesquels les nœuds sont répartis en sous-groupes qui partagent des propriétés semblables. On appellera ces sous-groupes des populations.

Définition 2.6. Soit G un graphe aléatoire d'ordre N . Une *division de G en populations* est une partition \mathcal{P} de $\{1, \dots, N\}$. On dit alors que chaque $J \in \mathcal{P}$ est une *population* de G . \diamond

L'intérêt de définir des populations sur un graphe aléatoire surviendra lorsque leurs nœuds partageront des propriétés similaires. C'est avec cette idée qu'on introduit la définition suivante.

Définition 2.7. Soient G un graphe aléatoire d'ordre N sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$ et \mathcal{P} une division de G en populations. Soit φ un paramètre d'ordre k sur G . On dit que φ est *homogène sur \mathcal{P}* ou *\mathcal{P} -homogène* si pour chaque choix $J_1, \dots, J_k \in \mathcal{P}$ de k populations, les variables aléatoires $\varphi_{j_1, \dots, j_k}$ pour $(j_1, \dots, j_k) \in J_1 \times \dots \times J_k$ sont indépendantes et identiquement distribuées. \diamond

On introduit également un cas particulier de paramètre homogène.

Définition 2.8. Soient G un graphe aléatoire d'ordre N sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$ et \mathcal{P} une division de G en populations. Soit φ un paramètre d'ordre k sur G qui est homogène sur \mathcal{P} . On dit que φ est *déterminé sur \mathcal{P}* ou *\mathcal{P} -déterminé* si pour chaque choix $J_1, \dots, J_k \in \mathcal{P}$ de k populations, il existe un $\varphi_{J_1, \dots, J_k} \in \mathbb{R}$ tel que $\varphi_{j_1, \dots, j_k} = \varphi_{J_1, \dots, J_k}$ μ -p.s. pour $(j_1, \dots, j_k) \in J_1 \times \dots \times J_k$. \diamond

§ 2.2. Dynamique markovienne ternaire cyclique

§ 2.2.1. Intuition sur la dynamique

Soit G un graphe aléatoire d'ordre N sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$. On veut maintenant définir aux nœuds d'un tel graphe aléatoire un état, et décrire l'évolution des états de ces nœuds selon un processus stochastique. L'idée est que ce processus attribue à chaque

nœud du graphe trois états possibles, et que ceux-ci évoluent de façon cyclique. On introduit la terminologie suivante pour décrire de tels états.

Définition 2.9. Soient $E := \{0, 1, i\}^N$, où i est l'unité imaginaire, et $\mathcal{E} := 2^E$. On dit que l'espace mesurable (E, \mathcal{E}) est un *espace d'états ternaires* pour un graphe aléatoire d'ordre N . On interprète alors un état $x = (x_j)_{j=1}^N \in E$ comme attribuant l'état x_j au nœud j du graphe. \diamond

Remarque 1. L'utilisation de l'ensemble $\{0, 1, i\}$ pour les valeurs possibles des états d'un nœud est nouvelle. Ce choix est fait afin que, si Z est une variable aléatoire qui prend ses valeurs dans $\{0, 1, i\}$, on peut voir les probabilités que Z soit dans un état donné comme des fonctions de Z . Par exemple, $\text{Re } Z$ correspond à la probabilité que $Z = 1$. Cette propriété sera cruciale dans les calculs qui surviendront à la section 2.3.

Remarque 2. Pour définir un processus semblable avec k états possibles, on pourrait choisir comme valeurs pour les états d'un nœud la base canonique $\{e_j\}_{j=1}^{n-1}$ de \mathbb{R}^{k-1} ainsi que le vecteur nul. Si Z est une variable aléatoire qui prend ses valeurs dans un tel ensemble, la projection de Z sur l'axe j correspond alors à la probabilité que $Z = e_j$. Le choix fait à la définition 2.9 est simplement le cas particulier de cette idée où $n = 2$, en identifiant $\mathbb{R}^2 = \mathbb{C}$.

On veut ensuite définir un processus de Markov en temps continu qui prendra ses valeurs dans E , de façon à attribuer au graphe un état à chaque instant. On voudra que les nœuds passent d'un état à l'autre selon les transitions $0 \mapsto 1 \mapsto i \mapsto 0$. Ces transitions seront appelées les transitions *permises*, alors que les autres transitions : $0 \mapsto 1$, $1 \mapsto i$ et $i \mapsto 0$, seront dites *interdites*.

Pour chaque nœud j du graphe, on caractérise ensuite les transitions permises par des fonctions $a_j, b_j, c_j: \Theta \times E \rightarrow [0, \infty)$, qui joueront le rôle de taux de transition. On exige que pour chaque état $x \in E$, les fonctions $\vartheta \mapsto a_j(\vartheta, x)$, $\vartheta \mapsto b_j(\vartheta, x)$ et $\vartheta \mapsto c_j(\vartheta, x)$ soient mesurables. Chaque taux de transition a_j, b_j ou c_j peut donc être vu comme une variable aléatoire sur $(\Theta, \mathcal{G}, \mu)$ pour chaque état $x \in E$. Par exemple, la dépendance des taux de transition sur Θ pourrait être donnée par une certaine fonction de paramètres sur le graphe, c'est-à-dire qu'un taux a_j pourrait avoir la forme $a_j(\vartheta, x) = \tilde{a}_j(\varphi_1(\vartheta), \varphi_2(\vartheta), \dots, \varphi_n(\vartheta), x)$ où $\tilde{a}_j: \mathbb{R}^d \times E \rightarrow [0, \infty)$ pour un certain $d \in \mathbb{N}$ et où les φ_ℓ seraient des paramètres sur le graphe.

Pour rendre plus explicite l'objectif d'introduire de tels taux de transition, on introduit $(\Omega, \mathcal{F}) := (E, \mathcal{E})^{[0, \infty)}$ et $\{X_t\}_{t \geq 0}$ le processus de projections de coordonnées sur Ω . Intuitivement, pour chaque $\vartheta \in \Theta$, on cherche à définir une mesure de probabilité \mathbb{P}^ϑ sur (Ω, \mathcal{F})

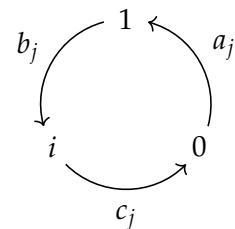


FIGURE 2.1 – Transitions permises entre les trois états.

telle que pour chaque j , lorsque Δt est assez petit,

$$a_j(\vartheta, x)\Delta t \approx \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1 \mid X_t^j = 0, X_t^j = x], \quad (2.1a)$$

$$b_j(\vartheta, x)\Delta t \approx \mathbb{P}^\vartheta[X_{t+\Delta t}^j = i \mid X_t^j = 1, X_t^j = x], \quad (2.1b)$$

$$c_j(\vartheta, x)\Delta t \approx \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 0 \mid X_t^j = i, X_t^j = x] \quad (2.1c)$$

et

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = i \mid X_t^j = 0] \approx \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 0 \mid X_t^j = 1] \approx \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1 \mid X_t^j = i] \approx 0 \quad (2.1d)$$

où X_t^j note la j -ième composante de X_t . Les taux de transition a_j , b_j et c_j décrivent en fait le générateur d'un processus de Markov. Si ce générateur est correctement défini, le théorème d'existence de Kolmogorov permettra d'introduire les mesures de probabilité voulues.

§ 2.2.2. Construction du processus de Markov

Soit $\vartheta \in \Theta$. Pour deux états $x = (x_j)_{j=1}^N$ et $y = (y_j)_{j=1}^N$ dans E , on définit

$$q^\vartheta(x, y) := \sum_{j=1}^N q_j^\vartheta(x, y) \prod_{\substack{k=1 \\ k \neq j}}^N \delta_{x_k y_k} \quad (2.2a)$$

où $\delta_{x_k y_k}$ est un delta de Kronecker et où

$$\begin{aligned} q_j^\vartheta(x, y) := & a_j(\vartheta, x)(1 - |x_j|)(\operatorname{Re} y_j - (1 - |y_j|)) + b_j(\vartheta, x) \operatorname{Re}(x_j)(\operatorname{Im} y_j - \operatorname{Re} y_j) \\ & + c_j(\vartheta, x) \operatorname{Im}(x_j)((1 - |y_j|) - \operatorname{Im} y_j). \end{aligned} \quad (2.2b)$$

On interprète $q^\vartheta(x, y)$ comme le taux de transition de l'état x vers l'état y . On remarque d'ailleurs que pour $x_j \in \{0, 1, i\}$, les trois expressions $1 - |x_j|$, $\operatorname{Re} x_j$ et $\operatorname{Im} x_j$ valent 1 si $x_j = 0$, si $x_j = 1$ et si $x_j = i$ respectivement, et sont nulles dans les autres cas. Ainsi, si la transition $x_j \mapsto y_j$ est permise, $q_j^\vartheta(x, y)$ correspond à $a_j(\vartheta, x)$, $b_j(\vartheta, x)$ ou $c_j(\vartheta, x)$ selon le cas, alors que $q_j^\vartheta(x, y) = 0$ si la transition $x_j \mapsto y_j$ est interdite.

On peut montrer que les q^ϑ forment effectivement le générateur d'une chaîne de Markov.

Lemme 2.2. Soit (E, \mathcal{E}) un espace d'états ternaires pour un graphe aléatoire, et soit la matrice $Q^\vartheta := \{q^\vartheta(x, y) : x, y \in E\}$ où $q^\vartheta(x, y)$ est défini tel qu'en (2.2). Alors :

- i) Pour $x, y \in E$ distincts, $q^\vartheta(x, y) \geq 0$;
- ii) Pour tout $x \in E$, $q^\vartheta(x, x) \leq 0$;
- iii) Pour tout $x \in E$, $\sum_{y \in E} q^\vartheta(x, y) = 0$.

◆

Démonstration. Soient $x, y \in E$ distincts. Si x et y diffèrent par au moins deux composantes, alors le produit $\prod_{k=1, k \neq j}^N \delta_{x_k y_k}$ est nul pour tout j et $q^\vartheta(x, y) = 0$. Sinon, x et y ne diffèrent que d'une composante j , et $q^\vartheta(x, y) = q_j^\vartheta(x, y)$. Comme $x_j, y_j \in \{0, 1, i\}$, (2.2b) donne que

$$q_j^\vartheta(x, y) = a_j(\vartheta, x)(1 - |x_j|) \operatorname{Re}(y_j) + b_j(\vartheta, x) \operatorname{Re}(x_j) \operatorname{Im}(y_j) + c_j(\vartheta, x) \operatorname{Im}(x_j)(1 - |y_j|). \quad (2.3)$$

On déduit que $q_j^\vartheta(x, y) \geq 0$, car a_j , b_j et c_j sont non négatives. Ceci termine de montrer i).

Soit $x \in E$. Évidemment, $\delta_{x_k x_k} = 1$ pour tout k . Ainsi, $q^\vartheta(x, x) = \sum_{j=1}^N q_j^\vartheta(x, x)$ d'après la définition (2.2). Cependant, pour chaque j , on a toujours que $x_j \in \{0, 1, i\}$, de sorte que d'après (2.2b),

$$q_j^\vartheta(x, x) = -a_j(\vartheta, x)(1 - |x_j|) - b_j(\vartheta, x) \operatorname{Re} x_j - c_j(\vartheta, x) \operatorname{Im} x_j. \quad (2.4)$$

On déduit alors que $q^\vartheta(x, x) = \sum_{j=1}^N q_j^\vartheta(x, x) \leq 0$, puisque les fonctions a_j , b_j et c_j sont non négatives. Ceci termine de montrer *ii*).

Soit $x \in E$. On peut écrire $\sum_{y \in E} q^\vartheta(x, y) = q^\vartheta(x, x) + \sum_{y \in E \setminus \{x\}} q^\vartheta(x, y)$. On a déjà vu que si x et y diffèrent par au moins deux composantes, alors $q^\vartheta(x, y) = 0$. De plus, si x et y ne diffèrent que par la j -ième composante, alors $q^\vartheta(x, y) = q_j^\vartheta(x, y)$. Dans ce cas, si la transition $x_j \mapsto y_j$ est interdite, alors $q^\vartheta(x, y) = 0$ puisque $q_j^\vartheta(x, y) = 0$. Tous les termes non nuls de la somme $\sum_{y \in E \setminus \{x\}} q^\vartheta(x, y)$ sont donc ceux où y ne diffère de x que d'une composante j et où la transition $x_j \mapsto y_j$ est permise. Pour chaque j , il y a donc un unique $y \in E \setminus \{x\}$ tel que $q^\vartheta(x, y)$ est non nul; on note \tilde{x}^j cet état. Alors $\sum_{y \in E \setminus \{x\}} q^\vartheta(x, y) = \sum_{j=1}^N q^\vartheta(x, \tilde{x}^j) = \sum_{j=1}^N q_j^\vartheta(x, \tilde{x}^j)$.

Pour chaque j , on a vu en (2.3) que $q_j^\vartheta(x, \tilde{x}^j)$ correspond à $a_j(\vartheta, x)$, $b_j(\vartheta, x)$ ou bien $c_j(\vartheta, x)$, selon la transition $x_j \mapsto \tilde{x}_j^j$. Cependant, d'après (2.4), $q_j^\vartheta(x, x)$ correspond justement au négatif de cette même fonction. Ainsi, pour chaque j on voit que $q_j^\vartheta(x, \tilde{x}_j^j) + q_j^\vartheta(x, x) = 0$. On obtient donc que

$$\sum_{y \in E} q^\vartheta(x, y) = q^\vartheta(x, x) + \sum_{y \in E \setminus \{x\}} q^\vartheta(x, y) = \sum_{j=1}^N (q_j^\vartheta(x, x) + q_j^\vartheta(x, \tilde{x}_j^j)) = 0. \quad (2.5)$$

Ceci termine de montrer *iii*). ■

Les propriétés de Q^ϑ obtenues au lemme précédent montrent qu'il s'agit en fait du générateur d'une chaîne de Markov à temps continu. Ainsi, pour chaque $\vartheta \in \Theta$, étant donnée une mesure de probabilité initiale sur (E, \mathcal{E}) , une mesure de probabilité peut être définie sur $(E, \mathcal{E})^{[0, \infty)}$ à partir du générateur Q^ϑ . Ces mesures peuvent même être reliées entre elles, comme montré au théorème suivant.

Théorème 2.3. Soit G un graphe aléatoire d'ordre N sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$. Soient (E, \mathcal{E}) un espace d'états ternaires sur G et P une mesure de probabilité sur (E, \mathcal{E}) . Pour chaque $j \in \{1, \dots, N\}$, soient $a_j, b_j, c_j: \Theta \times E \rightarrow [0, \infty)$ telles que pour tout $x \in E$, les projections $a_j(\cdot, x)$, $b_j(\cdot, x)$ et $c_j(\cdot, x)$ sont mesurables. Pour $\vartheta \in \Theta$, soit $q^\vartheta: E \times E \rightarrow \mathbb{R}$ définie par (2.2). Alors pour tout $\vartheta \in \Theta$, il existe une mesure de probabilité \mathbb{P}^ϑ sur $(\Omega, \mathcal{F}) := (E, \mathcal{E})^{[0, \infty)}$ telle que :

- i) Le processus de projection de coordonnées $\{X_t\}_{t \geq 0}$ sur $(\Omega, \mathcal{F}, \mathbb{P}^\vartheta)$ est un processus de Markov;
 - ii) Pour chaque $A \in \mathcal{E}$, $\mathbb{P}^\vartheta[X_0 \in A] = P(A)$;
 - iii) Pour $x, y \in E$, lorsque $\Delta t \downarrow 0$, $\mathbb{P}^\vartheta[X_{t+\Delta t} = y \mid X_t = x] = \delta_{xy} + q^\vartheta(x, y)\Delta t + o(\Delta t)$,
- et $v: (\vartheta, A) \mapsto \mathbb{P}^\vartheta(A)$ est une mesure de probabilité aléatoire de $(\Theta, \mathcal{G}, \mu)$ dans (Ω, \mathcal{F}) . ◆

Démonstration. Soit $\vartheta \in \Theta$. Par le lemme 2.2, la matrice $Q^\vartheta := \{q^\vartheta(x, y) : x, y \in E\}$ est le générateur d'une chaîne de Markov à temps continu homogène (voir définition 1.37). Ainsi, d'après la démarche détaillée à la section 1.2.5, le théorème d'extension de Kolmogorov montre l'existence d'une mesure de probabilité \mathbb{P}^ϑ sur (Ω, \mathcal{F}) qui respecte les conditions *i*), *ii*) et *iii*).

Il reste à montrer que $\nu : (\vartheta, A) \mapsto \mathbb{P}^\vartheta(A)$ définit bien une mesure de probabilité aléatoire de $(\Theta, \mathcal{G}, \mu)$ dans (Ω, \mathcal{F}) . On sait déjà que chaque \mathbb{P}^ϑ est une mesure de probabilité sur (Ω, \mathcal{F}) , donc il ne reste qu'à montrer que pour chaque $A \in \mathcal{F}$, la fonction $\vartheta \mapsto \mathbb{P}^\vartheta(A)$ est \mathcal{G} -mesurable.

On peut d'abord montrer que si $A = X_t^{-1}(B)$ pour un $B \in \mathcal{E}$ et un $t \geq 0$, alors la fonction $\vartheta \mapsto \mathbb{P}^\vartheta(A)$ est \mathcal{G} -mesurable. Pour chaque $\vartheta \in \Theta$, notons $p_t^\vartheta(x, y)$ les éléments de la matrice e^{tQ^ϑ} . D'après les résultats de la section 1.2.5, on sait que

$$\mathbb{P}^\vartheta(A) = \sum_{x_0 \in E} \sum_{x_1 \in B} P(\{x_0\}) p_t^\vartheta(x_0, x_1). \quad (2.6)$$

Or, d'après les propriétés des taux de transition a_j , b_j et c_j , les fonctions $\vartheta \mapsto q^\vartheta(x, y)$ sont \mathcal{G} -mesurables. Puisque l'exponentielle est une fonction mesurable, on déduit que les fonctions $\vartheta \mapsto p_t^\vartheta(x, y)$ sont aussi mesurables. L'expression (2.6) montre ainsi finalement que la fonction $\vartheta \mapsto \mathbb{P}^\vartheta(A)$ est \mathcal{G} -mesurable.

On définit ensuite $\mathcal{F}' := \{A \in \mathcal{F} : \vartheta \mapsto \mathbb{P}^\vartheta(A)$ est \mathcal{G} -mesurable}. Le raisonnement précédent montre que $\mathcal{F}' \supset \mathcal{B} := \{X_t^{-1}(B) : B \in \mathcal{E}, t \geq 0\}$. Or, on a vu à l'exemple 1.4 que la σ -algèbre produit $\mathcal{F} = \mathcal{E}^{[0, \infty)}$ est engendrée par le processus de projection de coordonnées, c'est-à-dire que $\mathcal{F} = \sigma(\mathcal{B})$. Ainsi, si on montre que \mathcal{F}' est une σ -algèbre, on aura montré que $\mathcal{F}' \supset \mathcal{F}$.

Puisque chaque \mathbb{P}^ϑ est une mesure de probabilité sur (Ω, \mathcal{F}) , on a toujours que $\mathbb{P}^\vartheta(\Omega) = 1$, donc la fonction $\vartheta \mapsto \mathbb{P}^\vartheta(\Omega)$ est \mathcal{G} -mesurable et $\Omega \in \mathcal{F}'$. Ensuite, si $A \in \mathcal{F}'$, alors $\mathbb{P}^\vartheta(\Omega \setminus A) = 1 - \mathbb{P}^\vartheta(A)$ par les propriétés des mesures de probabilité. Ainsi, comme la fonction $\vartheta \mapsto \mathbb{P}^\vartheta(A)$ est \mathcal{G} -mesurable, alors la fonction $\vartheta \mapsto \mathbb{P}^\vartheta(\Omega \setminus A) = 1 - \mathbb{P}^\vartheta(A)$ est \mathcal{G} -mesurable et $\Omega \setminus A \in \mathcal{F}'$. Finalement, soit $(A_k)_{k=1}^\infty \subset \mathcal{F}' \subset \mathcal{F}$ une suite d'ensembles disjoints. Comme \mathcal{F} est une σ -algèbre, la réunion $\bigcup_{k=1}^\infty A_k \in \mathcal{F}$. Ceci garantit que pour chaque $\vartheta \in \Theta$, $\sum_{k=1}^\infty \mathbb{P}^\vartheta(A_k) = \mathbb{P}^\vartheta(\bigcup_{k=1}^\infty A_k) \leq 1$, donc que cette série converge. Ainsi, puisque chaque fonction $\vartheta \mapsto \mathbb{P}^\vartheta(A_k)$ est \mathcal{G} -mesurable, la fonction $\vartheta \mapsto \mathbb{P}^\vartheta(\bigcup_{k=1}^\infty A_k) = \sum_{k=1}^\infty \mathbb{P}^\vartheta(A_k)$ est aussi \mathcal{G} -mesurable, et $\bigcup_{k=1}^\infty A_k \in \mathcal{F}'$.

\mathcal{F}' est donc une σ -algèbre. Puisque $\mathcal{F}' \supset \mathcal{B}$ et que $\mathcal{F} = \sigma(\mathcal{B})$, alors $\mathcal{F}' \supset \mathcal{F}$. Comme l'inclusion inverse est aussi vraie, on conclut alors que $\mathcal{F}' = \mathcal{F}$, c'est-à-dire que pour tout $A \in \mathcal{F}$, la fonction $\vartheta \mapsto \mathbb{P}^\vartheta(A)$ est \mathcal{G} -mesurable. La fonction ν est donc bien une mesure de probabilité aléatoire. ■

La propriété *iii)* que possèdent toutes les mesures \mathbb{P}^ϑ d'après le théorème 2.3 fournit de nombreuses propriétés. Le corollaire suivant en explicite quelques-unes qui seront particulièrement utiles plus loin.

Corollaire 1. Pour tout $\vartheta \in \Theta$ et tout $t \geq 0$, lorsque $\Delta t \downarrow 0$,

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1 \mid X_t^j = 0, X_t = x] = a_j(\vartheta, x)\Delta t + o(\Delta t), \quad (2.7a)$$

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = i \mid X_t^j = 1, X_t = x] = b_j(\vartheta, x)\Delta t + o(\Delta t), \quad (2.7b)$$

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = 0 \mid X_t^j = i, X_t = x] = c_j(\vartheta, x)\Delta t + o(\Delta t), \quad (2.7c)$$

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = 0 \mid X_t^j = 0, X_t = x] = 1 - a_j(\vartheta, x)\Delta t + o(\Delta t), \quad (2.7d)$$

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1 \mid X_t^j = 1, X_t = x] = 1 - b_j(\vartheta, x)\Delta t + o(\Delta t), \quad (2.7e)$$

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = i \mid X_t^j = i, X_t = x] = 1 - c_j(\vartheta, x)\Delta t + o(\Delta t), \quad (2.7f)$$

et que

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = x'_j \mid X_t^j = x_j, X_t = x] = o(\Delta t) \quad (2.7g)$$

lorsque la transition $x_j \mapsto x'_j$ est interdite. \blacklozenge

Démonstration. Ces propriétés sont toutes des conséquences de la propriété *iii)* des \mathbb{P}^ϑ donnée au théorème 2.3. En effet, prenons un $x \in E$ et un $\xi \in \{0, 1, i\}$ tel que $\xi \neq x_j$. Alors pour $t \geq 0$ et $\Delta t \downarrow 0$,

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = \xi \mid X_t = x] = \sum_{y \in E, y_j = \xi} \mathbb{P}^\vartheta[X_{t+\Delta t}^j = y \mid X_t = x] = \sum_{y \in E, y_j = \xi} (\delta_{xy} + q^\vartheta(x, y)\Delta t + o(\Delta t)). \quad (2.8)$$

On a ici toujours que $\delta_{xy} = 0$ puisque $\xi \neq x_j$. De plus, on a déjà vu que la définition (2.2) implique que $q^\vartheta(x, y) = 0$ lorsque y diffère de x d'au moins deux composantes. Ainsi, le seul état $y \in E$ avec $y_j = \xi$ tel que $q^\vartheta(x, y) \neq 0$ est celui qui ne diffère de x que par la j -ième composante. Notant z cet état, on obtient que

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = \xi \mid X_t = x] = q^\vartheta(x, z)\Delta t + o(\Delta t). \quad (2.9)$$

Si la transition $x_j \mapsto \xi$ est interdite, alors $q^\vartheta(x, z) = 0$ et $\mathbb{P}^\vartheta[X_{t+\Delta t}^j = \xi \mid X_t = x] = o(\Delta t)$, ce qui donne la relation (2.7g). Si la transition $x_j \mapsto \xi$ est permise, alors $q^\vartheta(x, z)$ vaut $a_j(\vartheta, x)$, $b_j(\vartheta, x)$ ou $c_j(\vartheta, x)$ selon la valeur de x_j , ce qui mène aux relations (2.7a), (2.7b) et (2.7c).

Finalement, on remarque que

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = x_j \mid X_t = x] = 1 - \mathbb{P}^\vartheta[X_{t+\Delta t}^j \neq x_j \mid X_t = x]. \quad (2.10)$$

Or, $X_{t+\Delta t}^j$ ne peut prendre que deux valeurs distinctes de x_j : une valeur ξ_1 telle que la transition $x_j \mapsto \xi_1$ est permise, et une autre valeur ξ_2 telle que la transition $x_j \mapsto \xi_2$ est interdite. Ainsi,

$$\mathbb{P}^\vartheta[X_{t+\Delta t}^j = x_j \mid X_t = x] = 1 - \mathbb{P}^\vartheta[X_{t+\Delta t}^j = \xi_1 \mid X_t = x] - \mathbb{P}^\vartheta[X_{t+\Delta t}^j = \xi_2 \mid X_t = x], \quad (2.11)$$

et les résultats (2.7d), (2.7e) et (2.7f) suivent à partir des premiers résultats obtenus. \blacksquare

Le fait de relier l'espace de probabilité $(\Theta, \mathcal{G}, \mu)$ à l'espace mesurable (Ω, \mathcal{F}) par une mesure de probabilité aléatoire permet de définir une mesure de probabilité sur le produit de ces deux espaces. Ceci donne une façon cohérente de voir le graphe aléatoire G et ses paramètres sur le même espace que le processus stochastique $\{X_t\}_{t \geq 0}$ qui décrit sa dynamique.

Corollaire 2. *Il existe une mesure de probabilité \mathbb{P} sur l'espace produit $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F})$ telle que si Z est une variable aléatoire sur $\Theta \times \Omega$,*

$$\mathbb{E}[Z] = \int_{\Theta} \int_{\Omega} Z(\vartheta, \omega) d\mathbb{P}^{\vartheta}(\omega) d\mu(\vartheta). \quad (2.12)$$

De plus, \mathbb{P} possède les propriétés suivantes :

- i) Si $A \in \mathcal{G}$, alors $\mathbb{P}(A \times \Omega) = \mu(A)$.
- ii) Si $B \in \mathcal{F}$, alors $\mathbb{P}(\Theta \times B) = \int_{\Theta} \mathbb{P}^{\vartheta}(B) d\mu(\vartheta)$ et $\mathbb{P}^{\vartheta}(B) = \mathbb{P}[\Theta \times B | \vartheta]$. \blacklozenge

Démonstration. Puisque $\nu: (\vartheta, A) \mapsto \mathbb{P}^{\vartheta}(A)$ est une mesure de probabilité aléatoire de $(\Theta, \mathcal{G}, \mu)$ dans (Ω, \mathcal{F}) , il suit directement du théorème 1.12 qu'il existe une mesure de probabilité \mathbb{P} sur $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F})$ qui respecte (2.12).

Les propriétés i) et ii) de \mathbb{P} sont simplement les résultats de calculs directs. En effet, si $A \in \mathcal{G}$, alors $\mathbb{P}(A \times \Omega) = \int_{\Theta} \int_{\Omega} \mathbb{1}_A(\vartheta) d\mathbb{P}^{\vartheta}(\omega) d\mu(\vartheta) = \mu(A)$. Ensuite, si $B \in \mathcal{F}$, alors $\mathbb{P}(\Theta \times B) = \int_{\Theta} \int_{\Omega} \mathbb{1}_B(\omega) d\mathbb{P}^{\vartheta}(\omega) d\mu(\vartheta) = \int_{\Theta} \mathbb{P}^{\vartheta}(B) d\mu(\vartheta)$. De plus, $\mathbb{P}^{\vartheta}(B) = \int_B d\mathbb{P}^{\vartheta} = \int_{\Omega} \mathbb{1}_{\Theta \times B}(\vartheta, \omega) d\mathbb{P}^{\vartheta}(\omega)$. Mais d'après la proposition 1.13, il s'agit de l'espérance conditionnelle $\mathbb{E}[\mathbb{1}_{\Theta \times B} | \vartheta] = \mathbb{P}[\Theta \times B | \vartheta]$. \blacksquare

On peut ici faire une remarque importante sur la notation. Étant donné un graphe aléatoire sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$, on a d'abord défini des paramètres sur le graphe comme des variables aléatoires sur $(\Theta, \mathcal{G}, \mu)$. On a ensuite construit un processus stochastique sur un espace mesurable (Ω, \mathcal{F}) afin d'attribuer au nœuds du graphe un état et de décrire leur évolution. Or, on voit finalement que l'on peut unifier ces deux espaces en un seul espace $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F}, \mathbb{P})$. Toutefois, n'importe quelle fonction mesurable sur (Θ, \mathcal{G}) ou sur (Ω, \mathcal{F}) peut naturellement être vue comme une variable aléatoire sur $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F}, \mathbb{P})$ en l'identifiant à une fonction qui ne dépend que d'une coordonnée. Comme il est plus simple de travailler sur un seul espace, dans tout ce qui suit, on fera de telles identifications sans toujours les mentionner explicitement, à moins qu'il y ait ambiguïté.

On introduit finalement la terminologie suivante pour référer aux structures décrites par le théorème.

Définition 2.10. Soit G un graphe aléatoire sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$. Si l'on décrit les états des nœuds de G par le processus stochastique $\{X_t\}_{t \geq 0}$ sur l'espace de probabilité $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F}, \mathbb{P})$ décrit au théorème 2.3 et ses corollaires, on dit que ce processus est une dynamique *markovienne ternaire cyclique* (MTC) sur G , décrite par les taux de transition a_j , b_j et c_j . Dans ce cas, on appelle $(\Theta, \mathcal{G}, \mu)$ l'espace des paramètres de G . \diamond

§ 2.3. Réduction à un système d'équations différentielles ordinaires

Soit G un graphe aléatoire d'ordre N sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$. Soit un processus stochastique $\{X_t\}_{t \geq 0}$ qui définit sur G une dynamique MTC décrite par les taux a_j , b_j et c_j sur l'espace $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F}, \mathbb{P})$. Bien que ce processus stochastique décrive complètement la dynamique de G , il ne fournit pas une description idéale du système, car il n'est pas évident de l'étudier lorsque N est grand. Or, dans les applications, on s'intéresse généralement à cette situation. De plus, les taux de transition a_j , b_j et c_j dépendent en général d'un choix de $\vartheta \in \Theta$. Dans les applications, ces taux sont donnés par des fonctions de paramètres sur le graphe ; la dépendance des taux sur ϑ correspond alors à une dépendance sur les valeurs de chaque paramètre en chaque nœud du graphe. Or, dans ces situations, il est plutôt utopique de chercher à déterminer les valeurs des paramètres en chaque nœud. L'information dont on dispose peut plutôt correspondre à une loi de probabilité qui modéliseraient les valeurs des paramètres, par exemple.

Étant données ces difficultés, on cherche alors à tirer du processus stochastique une autre description de la dynamique définie sur G , qui décrirait l'évolution d'un plus petit nombre de variables que le processus stochastique, et qui ne dépendrait pas directement d'un choix $\vartheta \in \Theta$. Il est bien connu (voir par exemple [66] ou plus récemment [23]) que certains processus de Markov ont des limites sous la forme de solutions d'équations différentielles ordinaires. Il est donc naturel de chercher à approximer le processus de Markov décrit à la section précédente par la solution d'un tel système. On cherche donc dans un premier temps à décrire les dérivées des probabilités que chaque nœud soit dans un état donné.

Pour chaque nœud j et chaque $\vartheta \in \Theta$, on introduit les fonctions $p_j^\vartheta, q_j^\vartheta, r_j^\vartheta: [0, \infty) \rightarrow [0, 1]$ données par

$$p_j^\vartheta(t) := \mathbb{P}[X_t^j = 1 \mid \vartheta], \quad q_j^\vartheta(t) := \mathbb{P}[X_t^j = 0 \mid \vartheta] \quad \text{et} \quad r_j^\vartheta(t) := \mathbb{P}[X_t^j = i \mid \vartheta], \quad (2.13)$$

ainsi que les fonctions $p_j, q_j, r_j: [0, \infty) \rightarrow [0, 1]$ données par

$$p_j(t) := \mathbb{P}[X_t^j = 1], \quad q_j(t) := \mathbb{P}[X_t^j = 0] \quad \text{et} \quad r_j(t) := \mathbb{P}[X_t^j = i]. \quad (2.14)$$

On remarque que, puisque les X_t sont définis sur Ω , il suit de la propriété *ii)* de \mathbb{P} donnée au corollaire 2 du théorème 2.3 que $p_j(t) = \int_{\Theta} p_j^\vartheta(t) d\mu(\vartheta)$, et de même pour q_j et r_j .

Afin de manipuler plus facilement ces probabilités, on obtient d'abord le résultat suivant.

Lemme 2.4. *Soit G un graphe aléatoire muni d'une dynamique MTC. Pour chaque nœud j de G , soient $p_j^\vartheta, q_j^\vartheta, r_j^\vartheta: [0, \infty) \rightarrow [0, 1]$ définies par (2.13). Alors $p_j^\vartheta + q_j^\vartheta + r_j^\vartheta \equiv 1$, et pour $t \geq 0$ et $\vartheta \in \Theta$,*

$$p_j^\vartheta(t) = \mathbb{E}[\operatorname{Re} X_t^j \mid \vartheta], \quad q_j^\vartheta(t) = \mathbb{E}[1 - |X_t^j| \mid \vartheta] \quad \text{et} \quad r_j^\vartheta(t) = \mathbb{E}[\operatorname{Im} X_t^j \mid \vartheta]. \quad (2.15)$$

Démonstration. Soient $t \geq 0$, j un nœud de G et $\vartheta \in \Theta$. On sait que X_t^j prend ses valeurs dans $\{0, 1, i\}$, donc puisque $A \mapsto \mathbb{P}[A | \vartheta]$ est une mesure de probabilité sur $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F})$,

$$p_j^\vartheta(t) + q_j^\vartheta(t) + r_j^\vartheta(t) = \mathbb{P}[X_t^j = 1 | \vartheta] + \mathbb{P}[X_t^j = 0 | \vartheta] + \mathbb{P}[X_t^j = i | \vartheta] = 1. \quad (2.16)$$

Ensuite,

$$p_j^\vartheta(t) = \mathbb{P}[X_t^j = 1 | \vartheta] = \text{Re}(1)\mathbb{P}[X_t^j = 1 | \vartheta] = \sum_{x \in \{0, 1, i\}} \text{Re}(x)\mathbb{P}[X_t^j = x | \vartheta] = \mathbb{E}[\text{Re } X_t^j | \vartheta], \quad (2.17)$$

et de la même façon

$$r_j^\vartheta(t) = \mathbb{P}[X_t^j = i | \vartheta] = \text{Im}(i)\mathbb{P}[X_t^j = i | \vartheta] = \sum_{x \in \{0, 1, i\}} \text{Im}(x)\mathbb{P}[X_t^j = x | \vartheta] = \mathbb{E}[\text{Im } X_t^j | \vartheta]. \quad (2.18)$$

Les résultats précédents impliquent alors que

$$q_j^\vartheta(t) = 1 - p_j^\vartheta(t) - r_j^\vartheta(t) = 1 - \mathbb{E}[\text{Re } X_t^j | \vartheta] - \mathbb{E}[\text{Im } X_t^j | \vartheta] = \mathbb{E}[1 - \text{Re } X_t^j - \text{Im } X_t^j | \vartheta]. \quad (2.19)$$

Or, $\text{Re } X_t^j + \text{Im } X_t^j$ vaut 1 si $X_t^j = 1$ ou $X_t^j = i$ et vaut 0 si $X_t^j = 0$. Comme ces trois cas sont les seuls possibles, on voit que $\text{Re } X_t^j + \text{Im } X_t^j = |X_t^j|$, de sorte que $q_j(t) = \mathbb{E}[1 - |X_t^j| | \vartheta]$. ■

On obtient ensuite directement l'analogue de ce résultat, mais sans conditionner sur la projection sur l'espace des paramètres.

Lemme 2.5. Soit G un graphe aléatoire muni d'une dynamique MTC. Pour chaque nœud j de G , soient $p_j, q_j, r_j: [0, \infty) \rightarrow [0, 1]$ définies par (2.14). Alors $p_j + q_j + r_j \equiv 1$, et pour $t \geq 0$,

$$p_j(t) = \mathbb{E}[\text{Re } X_t^j], \quad q_j(t) = \mathbb{E}[1 - |X_t^j|] \quad \text{et} \quad r_j(t) = \mathbb{E}[\text{Im } X_t^j]. \quad (2.20)$$

◆

Démonstration. Soient $t \geq 0$ et j un nœud de G . Par le lemme 2.4 on sait que pour tout $\vartheta \in \Theta$, $p_j^\vartheta + q_j^\vartheta + r_j^\vartheta \equiv 1$. Ainsi,

$$p_j(t) + q_j(t) + r_j(t) = \int_{\Theta} (p_j^\vartheta(t) + q_j^\vartheta(t) + r_j^\vartheta(t)) \, d\mu(\vartheta) = \int_{\Theta} d\mu = 1. \quad (2.21)$$

Ensuite, on a vu que $p_j^\vartheta(t) = \mathbb{E}[\text{Re } X_t^j | \vartheta]$. Ainsi, en utilisant la proposition 1.13 pour obtenir une forme intégrale de cette espérance conditionnelle, on trouve que

$$p_j(t) = \int_{\Theta} \mathbb{E}[\text{Re } X_t^j | \vartheta] \, d\mu(\vartheta) = \int_{\Theta} \int_{\Omega} \text{Re } X_t^j(\omega) \, d\mathbb{P}^\vartheta(\omega) \, d\mu(\vartheta) = \mathbb{E}[\text{Re } X_t^j]. \quad (2.22)$$

Les identités pour q_j et r_j sont obtenues de la même façon. ■

À partir de ce résultat et des propriétés des mesures \mathbb{P}^ϑ , on peut ensuite obtenir des expressions pour les dérivées des p_j^ϑ , des q_j^ϑ et des r_j^ϑ .

Théorème 2.6. Pour tout $\vartheta \in \Theta$, les fonctions p_j^ϑ , q_j^ϑ et r_j^ϑ introduites au lemme 2.4 sont dérивables sur $(0, \infty)$ et dérивables à droite en 0. De plus, pour tout $t \geq 0$,

$$\dot{p}_j^\vartheta(t) = -\mathbb{E}[\operatorname{Re}(X_t^j)b_j(\vartheta, X_t) | \vartheta] + \mathbb{E}[(1 - |X_t^j|)a_j(\vartheta, X_t) | \vartheta], \quad (2.23a)$$

$$\dot{q}_j^\vartheta(t) = -\mathbb{E}[(1 - |X_t^j|)a_j(\vartheta, X_t) | \vartheta] + \mathbb{E}[\operatorname{Im}(X_t^j)c_j(\vartheta, X_t) | \vartheta], \quad (2.23b)$$

$$\dot{r}_j^\vartheta(t) = -\mathbb{E}[\operatorname{Im}(X_t^j)c_j(\vartheta, X_t) | \vartheta] + \mathbb{E}[\operatorname{Re}(X_t^j)b_j(\vartheta, X_t) | \vartheta], \quad (2.23c)$$

où le point note la dérivée. \blacklozenge

Remarque. C'est ici que l'on voit toute la pertinence d'avoir choisi $\{0, 1, i\}$ comme l'ensemble des valeurs possibles pour les états d'un nœud du graphe. La méthode utilisée dans la preuve pour calculer les dérivées de ces probabilités est inspirée d'une méthode classique en épidémiologie, utilisée notamment dans [15, 94, 95]. Dans ce contexte, on considère un processus stochastique $\{Z_t\}_{t \geq 0}$ où chaque Z_t prend ses valeurs dans $\{0, 1\}$. La dérivée de la probabilité que $Z_t = 1$ est alors calculée en exploitant le fait que cette probabilité est égale à l'espérance de Z_t . Prendre $\{0, 1, i\}$ comme valeurs pour les états d'un nœud permet de généraliser cette méthode.

Démonstration. Soit $\vartheta \in \Theta$. Pour alléger la notation, on travaillera sur $(\Omega, \mathcal{F}, \mathbb{P}^\vartheta)$ directement afin d'éviter le conditionnement sur ϑ . Cependant, d'après la proposition 1.13, si Y est mesurable sur (Ω, \mathcal{F}) , alors $\mathbb{E}[Y | \vartheta] = \mathbb{E}^\vartheta[Y]$, où \mathbb{E}^ϑ note l'espérance sur $(\Omega, \mathcal{F}, \mathbb{P}^\vartheta)$. On démontrera donc l'équivalent des relations (2.23) sur $(\Omega, \mathcal{F}, \mathbb{P}^\vartheta)$.

Soient $t \geq 0$ et $\Delta t > 0$. On remarque que l'ensemble $\{X_{t+\Delta t}^j = 1\} \subset \Omega$ peut être écrit comme $\bigcup_{x \in E} \{X_t = x, X_{t+\Delta t}^j = 1\}$, qui est une réunion d'ensembles disjoints. Ainsi, pour $\vartheta \in \Theta$,

$$p_j^\vartheta(t + \Delta t) = \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1] = \mathbb{P}^\vartheta\left(\bigcup_{x \in E} \{X_{t+\Delta t}^j = 1, X_t = x\}\right) \quad (2.24)$$

$$= \sum_{x \in E} \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1, X_t = x] = \sum_{x \in E} \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1 | X_t = x] \mathbb{P}^\vartheta[X_t = x]. \quad (2.25)$$

On peut ensuite utiliser l'identité

$$\begin{aligned} \mathbb{P}^\vartheta[X_{t+\Delta t}^j = y | X_t = x] &= \operatorname{Re}(x_j) \mathbb{P}^\vartheta[X_{t+\Delta t}^j = y | X_t = x, X_t^j = 1] \\ &\quad + \operatorname{Im}(x_j) \mathbb{P}^\vartheta[X_{t+\Delta t}^j = y | X_t = x, X_t^j = i] \\ &\quad + (1 - |x_j|) \mathbb{P}^\vartheta[X_{t+\Delta t}^j = y | X_t = x, X_t^j = 0], \end{aligned} \quad (2.26)$$

pour $y \in \{0, 1, i\}$. Cette identité suit directement du fait que x_j doit prendre une valeur dans $\{0, 1, i\}$: deux des termes du membre de droite de (2.26) sont nuls, et l'autre correspond directement au membre de gauche en y explicitant la valeur de x_j dans la condition $X_t = x$. En utilisant (2.26) avec $y = 1$ dans (2.25), on obtient que

$$\begin{aligned} p_j^\vartheta(t + \Delta t) &= \sum_{x \in E} \left(\operatorname{Re}(x_j) \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1 | X_t = x, X_t^j = 1] \right. \\ &\quad \left. + \operatorname{Im}(x_j) \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1 | X_t = x, X_t^j = i] \right. \\ &\quad \left. + (1 - |x_j|) \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1 | X_t = x, X_t^j = 0] \right) \mathbb{P}^\vartheta[X_t = x]. \end{aligned} \quad (2.27)$$

En considérant $\Delta t \downarrow 0$, il est justifié d'utiliser les expressions (2.7) obtenues au corollaire 1 du théorème 2.3 pour les probabilités conditionnelles de l'expression précédente. On obtient alors que

$$\begin{aligned} p_j^\vartheta(t + \Delta t) &= \sum_{x \in E} \left(\operatorname{Re}(x_j) (1 - b_j(\vartheta, x) \Delta t + o(\Delta t)) + \operatorname{Im}(x_j) o(\Delta t) \right. \\ &\quad \left. + (1 - |x_j|) (a_j(\vartheta, x) \Delta t + o(\Delta t)) \right) \mathbb{P}^\vartheta[X_t = x]. \end{aligned} \quad (2.28)$$

$$= \sum_{x \in E} \left(\operatorname{Re}(x_j) - b_j(\vartheta, x) \operatorname{Re}(x_j) \Delta t + a_j(\vartheta, x) (1 - |x_j|) \Delta t \right) \mathbb{P}^\vartheta[X_t = x] + o(\Delta t). \quad (2.29)$$

Mais E est l'ensemble des valeurs possibles pour X_t . Ainsi,

$$p_j^\vartheta(t + \Delta t) = \mathbb{E}^\vartheta [\operatorname{Re}(X_t^j) - b_j(\vartheta, X_t) \operatorname{Re}(X_t^j) \Delta t + a_j(\vartheta, X_t) (1 - |X_t^j|) \Delta t] + o(\Delta t). \quad (2.30)$$

Par le lemme 2.4, on sait que $p_j^\vartheta(t) = \mathbb{E}^\vartheta[\operatorname{Re} X_t^j]$. On trouve alors finalement que

$$\frac{p_j^\vartheta(t + \Delta t) - p_j^\vartheta(t)}{\Delta t} = \mathbb{E}^\vartheta [-b_j(\vartheta, X_t) \operatorname{Re}(X_t^j) + a_j(\vartheta, X_t) (1 - |X_t^j|)] + \frac{1}{\Delta t} o(\Delta t). \quad (2.31)$$

Puisque $\frac{1}{\Delta t} o(\Delta t) \rightarrow 0$ lorsque $\Delta t \downarrow 0$, la relation précédente montre que la dérivée à droite de p_j^ϑ existe et est donnée par

$$\mathbb{E}^\vartheta [-b_j(\vartheta, X_t) \operatorname{Re}(X_t^j) + a_j(\vartheta, X_t) (1 - |X_t^j|)]. \quad (2.32)$$

Cependant, comme on a remarqué à la fin de la section 1.2.5, l'équation de Kolmogorov (1.30) implique que p_j^ϑ est dérivable sur $(0, \infty)$ puisque $\{X_t\}_{t \geq 0}$ est un processus de Markov. L'expression (2.32) donne donc la dérivée de p_j^ϑ sur $(0, \infty)$ et sa dérivée à droite en 0.

Un raisonnement très similaire au précédent mène à l'expression de la dérivée à droite de $r_j^\vartheta(t)$. En effet, de la même façon qu'on a obtenu (2.25), on peut voir que

$$r_j^\vartheta(t + \Delta t) = \sum_{x \in E} \mathbb{P}^\vartheta[X_{t+\Delta t}^j = i \mid X_t = x] \mathbb{P}^\vartheta[X_t = x]. \quad (2.33)$$

Ensuite, en utilisant (2.26) avec $y = i$ et en substituant les expressions (2.7), on trouve que

$$\begin{aligned} r_j^\vartheta(t + \Delta t) &= \sum_{x \in E} \left(\operatorname{Re}(x_j) \mathbb{P}^\vartheta[X_{t+\Delta t}^j = i \mid X_t = x, X_t^j = 1] \right. \\ &\quad \left. + \operatorname{Im}(x_j) \mathbb{P}^\vartheta[X_{t+\Delta t}^j = i \mid X_t = x, X_t^j = i] \right. \\ &\quad \left. + (1 - |x_j|) \mathbb{P}^\vartheta[X_{t+\Delta t}^j = i \mid X_t = x, X_t^j = 0] \right) \mathbb{P}^\vartheta[X_t = x] \end{aligned} \quad (2.34)$$

$$= \sum_{x \in E} \left(\operatorname{Re}(x_j) b_j(\vartheta, x) \Delta t + \operatorname{Im}(x_j) (1 - c_j(\vartheta, x) \Delta t) \right) \mathbb{P}^\vartheta[X_t = x] + o(\Delta t). \quad (2.35)$$

$$= \mathbb{E}^\vartheta [\operatorname{Re}(X_t^j) b_j(\vartheta, X_t) \Delta t + \operatorname{Im}(X_t^j) (1 - c_j(\vartheta, X_t) \Delta t)] + o(\Delta t). \quad (2.36)$$

Par le lemme 2.4, on sait que $r_j^\vartheta(t) = \mathbb{E}^\vartheta[\text{Im}(X_t^j)]$. Ainsi,

$$\frac{r_j^\vartheta(t + \Delta t) - r_j^\vartheta(t)}{\Delta t} = \mathbb{E}^\vartheta[\text{Re}(X_t^j)b_j(\vartheta, X_t) - \text{Im}(X_t^j)c_j(\vartheta, X_t)] + \frac{1}{\Delta t}o(\Delta t), \quad (2.37)$$

et on voit que la dérivée à droite de r_j^ϑ est donnée par

$$\mathbb{E}^\vartheta[\text{Re}(X_t^j)b_j(\vartheta, X_t) - \text{Im}(X_t^j)c_j(\vartheta, X_t)]. \quad (2.38)$$

Comme plus tôt, on sait d'après l'équation de Kolmogorov que r_j^ϑ est dérivable. Cette expression donne ainsi la dérivée de r_j^ϑ sur $(0, \infty)$.

Finalement, le lemme 2.4 montre que $p_j^\vartheta + q_j^\vartheta + r_j^\vartheta \equiv 1$. Ainsi, les résultats obtenus pour \dot{p}_j^ϑ et \dot{r}_j^ϑ montrent directement que

$$\dot{q}_j^\vartheta(t) = -\dot{p}_j^\vartheta(t) - \dot{r}_j^\vartheta(t) = -\mathbb{E}^\vartheta[(1 - |X_t^j|)a_j(\vartheta, X_t)] + \mathbb{E}^\vartheta[\text{Im}(X_t^j)c_j(\vartheta, X_t)]. \quad (2.39)$$

■

Les résultats du théorème (2.6) sont un premier pas vers une description travaillable de la dynamique sur le graphe. Cependant, comme décrit plus tôt, on cherche une description qui ne dépendrait pas du choix de paramètres, c'est-à-dire de la valeur précise de ϑ dans Θ . Heureusement, le résultat précédent a un analogue sur les probabilités p_j , q_j et r_j .

Corollaire. Si pour tout $x \in E$ et pour tout nœud j de G les fonctions $a_j(\cdot, x)$, $b_j(\cdot, x)$ et $c_j(\cdot, x)$ sont intégrables, alors les fonctions p_j , q_j et r_j introduites au lemme 2.5 sont dériviales sur $(0, \infty)$ et dériviales à droite en 0, et pour tout $t \geq 0$,

$$\dot{p}_j(t) = -\mathbb{E}[\text{Re}(X_t^j)b_j(\vartheta, X_t)] + \mathbb{E}[(1 - |X_t^j|)a_j(\vartheta, X_t)], \quad (2.40a)$$

$$\dot{q}_j(t) = -\mathbb{E}[(1 - |X_t^j|)a_j(\vartheta, X_t)] + \mathbb{E}[\text{Im}(X_t^j)c_j(\vartheta, X_t)], \quad (2.40b)$$

$$\dot{r}_j(t) = -\mathbb{E}[\text{Im}(X_t^j)c_j(\vartheta, X_t)] + \mathbb{E}[\text{Re}(X_t^j)b_j(\vartheta, X_t)]. \quad (2.40c)$$

◆

Démonstration. On note d'abord que, pour tout $\vartheta \in \Theta$ et pour tout $t \geq 0$, le théorème 2.6 montre que

$$|\dot{p}_j^\vartheta(t)| = \left| -\mathbb{E}[\text{Re}(X_t^j)b_j(\vartheta, X_t) \mid \vartheta] + \mathbb{E}[(1 - |X_t^j|)a_j(\vartheta, X_t) \mid \vartheta] \right| \quad (2.41)$$

$$\leq \mathbb{E}[|\text{Re}(X_t^j)b_j(\vartheta, X_t)| \mid \vartheta] + \mathbb{E}[|(1 - |X_t^j|)a_j(\vartheta, X_t)| \mid \vartheta] \quad (2.42)$$

$$\leq \mathbb{E}[|b_j(\vartheta, X_t)| + |a_j(\vartheta, X_t)| \mid \vartheta]. \quad (2.43)$$

Comme E est fini, on peut définir

$$f(\vartheta) := \sum_{x \in E} (|a_j(\vartheta, x)| + |b_j(\vartheta, x)|), \quad (2.44)$$

qui est intégrable puisque l'on suppose que les fonctions $a_j(\cdot, x)$ et $b_j(\cdot, x)$ sont intégrables

pour chaque $x \in E$. Cette fonction permet de borner $|p_j^\vartheta(t)|$: d'après (2.43),

$$|p_j^\vartheta(t)| \leq \sum_{x \in E} \mathbb{E}[|b_j(\vartheta, x)| + |a_j(\vartheta, x)| \mid \vartheta] = \mathbb{E}[f(\vartheta) \mid \vartheta] = f(\vartheta). \quad (2.45)$$

Cette condition de domination justifie alors de passer la dérivée sous l'intégrale, et on trouve que

$$\dot{p}_j(t) = \frac{d}{dt} \int_{\Theta} p_j^\vartheta(t) d\mu(\vartheta) = \int_{\Theta} \dot{p}_j^\vartheta(t) d\mu(\vartheta). \quad (2.46)$$

L'expression (2.40a) suit ensuite directement de l'intégration de l'expression de $\dot{p}_j^\vartheta(t)$ obtenue au théorème 2.6, en utilisant la proposition 1.13.

Des preuves analogues permettent de démontrer les résultats (2.40b) et (2.40c). ■

Le système d'équations différentielles obtenu au corollaire précédent donne une description de la dynamique qui ne dépend plus de la valeur précise des paramètres. Toutefois, pour pouvoir étudier ce système, il serait nécessaire de réécrire les membres de droite de (2.40) en termes des p_j , des q_j et des r_j . Or, il n'est pas clair que ce soit possible, à moins d'ajouter davantage de suppositions sur les taux de transition a_j , b_j et c_j . Par ailleurs, même si l'on pouvait obtenir un tel système, il resterait difficile à étudier, puisque si le nombre de nœuds du graphe est très grand, la dimension de ce système d'équations différentielles est encore plus grande.

§ 2.4. Description macroscopique de la dynamique

Le système (2.40) obtenu au corollaire du théorème 2.6, à la section précédente, fournit une description de la dynamique définie sur le graphe qui ne dépend plus du choix précis de paramètres sur le graphe. Cependant, ce système est d'une très grande dimension : il décrit l'évolution des états de tous les nœuds.

Pour obtenir une description de la dynamique qui soit plus simple à étudier, il est alors logique de s'intéresser à réduire le système (2.40) à une description macroscopique de la dynamique. Pour ce faire, on suppose que le graphe G est divisé en une collection \mathcal{P} de n populations. On peut alors chercher à décrire comment ces populations évoluent, plutôt que de chercher à décrire comment les nœuds évoluent individuellement.

Pour chaque population $J \in \mathcal{P}$ et chaque $t \geq 0$, on définit

$$A_t^J := \frac{1}{|J|} \sum_{j \in J} \operatorname{Re} X_t^j, \quad R_t^J := \frac{1}{|J|} \sum_{j \in J} \operatorname{Im} X_t^j, \quad \text{et} \quad S_t^J := \frac{1}{|J|} \sum_{j \in J} (1 - |X_t^j|), \quad (2.47)$$

qui correspondent aux fractions des nœuds de J qui se trouvent dans l'état 1, dans l'état i et dans l'état 0 respectivement. La paire (A_t^J, R_t^J) peut alors être interprétée comme l'état dans lequel se trouve la population J d'un point de vue macroscopique. On note que si l'on

connaît les fractions A_t^J et R_t^J pour une certaine population J , alors on connaît immédiatement S_t^J . En effet, comme chaque X_t^j prend ses valeurs dans $\{0, 1, i\}$, alors

$$A_t^J + R_t^J + S_t^J = \frac{1}{|J|} \sum_{j \in J} (\operatorname{Re} X_t^j + \operatorname{Im} X_t^j + 1 - |X_t^j|) \equiv 1. \quad (2.48)$$

En combinant les états de toutes les populations comme

$$Y_t := (Y_t^J)_{J \in \mathcal{P}} \quad \text{où} \quad Y_t^J := (A_t^J, R_t^J), \quad (2.49)$$

on obtient un nouveau processus stochastique $\{Y_t\}_{t \geq 0}$ qui décrit l'évolution de l'état du graphe d'un point de vue macroscopique.

Pour obtenir une description de l'évolution attendue de l'état du graphe, on peut alors s'intéresser aux espérances des composantes de Y_t . Afin d'alléger la notation pour la suite, on notera ces espérances par des lettres calligraphiques :

$$\mathcal{A}_J(t) := \mathbb{E}[A_t^J], \quad \mathcal{R}_J(t) := \mathbb{E}[R_t^J] \quad \text{et} \quad \mathcal{S}_J(t) := \mathbb{E}[S_t^J]. \quad (2.50)$$

Par linéarité de l'espérance, on remarque alors que

$$\mathcal{A}_J(t) = \mathbb{E}[A_t^J] = \mathbb{E}\left[\frac{1}{|J|} \sum_{j \in J} \operatorname{Re} X_t^j\right] = \frac{1}{|J|} \sum_{j \in J} p_j(t). \quad (2.51)$$

De la même façon,

$$\mathcal{R}_J(t) = \frac{1}{|J|} \sum_{j \in J} r_j(t) \quad \text{et} \quad \mathcal{S}_J(t) = \frac{1}{|J|} \sum_{j \in J} q_j(t). \quad (2.52)$$

Ainsi, sous les hypothèses du corollaire du théorème 2.6, la dérivabilité des p_j , des q_j et des r_j garantit la dérivabilité des \mathcal{A}_J , des \mathcal{R}_J et des \mathcal{S}_J . Le résultat (2.40) donne alors des expressions pour les dérivées de ces quantités macroscopiques, par linéarité de la dérivée.

À partir de cette idée, on peut obtenir une forme assez évocatrice pour les dérivées des \mathcal{A}_J , des \mathcal{R}_J et des \mathcal{S}_J . Pour ce faire, on définit les ensembles aléatoires

$$J_t^\xi := \{j \in J : X_t^j = \xi\} \quad (2.53)$$

pour $\xi \in \{0, 1, i\}$, et on remarque que

$$A_t^J = \frac{|J_t^1|}{|J|}, \quad R_t^J = \frac{|J_t^i|}{|J|} \quad \text{et} \quad S_t^J = \frac{|J_t^0|}{|J|}. \quad (2.54)$$

Ainsi, à partir de (2.40), on voit que

$$\dot{\mathcal{A}}_J(t) = \frac{1}{|J|} \sum_{j \in J} \dot{p}_j^\vartheta(t) = \frac{1}{|J|} \sum_{j \in J} \left(-\mathbb{E}[\operatorname{Re}(X_t^j) b_j(\vartheta, X_t)] + \mathbb{E}[(1 - |X_t^j|) a_j(\vartheta, X_t)] \right) \quad (2.55)$$

$$= \mathbb{E}\left[-\frac{|J_t^1|}{|J|} \frac{1}{|J_t^1|} \sum_{j \in J} \operatorname{Re}(X_t^j) b_j(\vartheta, X_t) + \frac{|J_t^0|}{|J|} \frac{1}{|J_t^0|} (1 - |X_t^j|) a_j(\vartheta, X_t) \right] \quad (2.56)$$

$$= \mathbb{E}\left[-\frac{A_t^J}{|J_t^1|} \sum_{j \in J} \operatorname{Re}(X_t^j) b_j(\vartheta, X_t) + \frac{S_t^J}{|J_t^0|} \sum_{j \in J} (1 - |X_t^j|) a_j(\vartheta, X_t) \right]. \quad (2.57)$$

Cependant, $j \in J_t^1$ si et seulement si $\text{Re}(X_t^j) = 1$, et si $j \notin J_t^1$ alors $\text{Re}(X_t^j) = 0$. De la même façon, $j \in J_t^0$ si et seulement si $1 - |X_t^j| = 1$, et $1 - |X_t^j| = 0$ lorsque $j \notin J_t^0$. Il s'ensuit que

$$\dot{\mathcal{A}}_J(t) = \mathbb{E} \left[-A_t^J \frac{1}{|J_t^1|} \sum_{j \in J_t^1} b_j(\vartheta, X_t) + S_t^J \frac{1}{|J_t^0|} \sum_{j \in J_t^0} a_j(\vartheta, X_t) \right]. \quad (2.58a)$$

Les mêmes idées mènent, dans les deux autres cas, aux résultats

$$\dot{\mathcal{R}}_J(t) = \mathbb{E} \left[-R_t^J \frac{1}{|J_t^1|} \sum_{j \in J_t^1} c_j(\vartheta, X_t) + A_t^J \frac{1}{|J_t^1|} \sum_{j \in J_t^1} b_j(\vartheta, X_t) \right] \quad (2.58b)$$

et

$$\dot{\mathcal{S}}_J(t) = \mathbb{E} \left[-S_t^J \frac{1}{|J_t^0|} \sum_{j \in J_t^0} a_j(\vartheta, X_t) + R_t^J \frac{1}{|J_t^0|} \sum_{j \in J_t^0} c_j(\vartheta, X_t) \right]. \quad (2.58c)$$

Ces résultats donnent une intuition claire sur la façon dont on peut décrire la dynamique sur le graphe d'un point de vue macroscopique. En effet, les taux de transition n'apparaissent pas individuellement dans ces équations : on n'en voit que des moyennes arithmétiques prises sur les fractions d'une population dans un état donné. Si ces moyennes arithmétiques peuvent être réécrites comme des taux de transition moyens qui ne dépendent que de l'état macroscopique du graphe, alors (2.58) permettra d'écrire les dérivées des \mathcal{A}_J , des \mathcal{R}_J et des \mathcal{S}_J en fonction de l'état macroscopique du graphe plutôt qu'on fonction de son état microscopique.

On suppose donc qu'il existe des fonctions $a_J, b_J, c_J: [0, 1]^{2n} \rightarrow [0, \infty)$, où n est le nombre de populations du graphe, telles que

$$a_J(Y_t) = \frac{1}{|J_t^0|} \sum_{j \in J_t^0} a_j(\vartheta, X_t), \quad b_J(Y_t) = \frac{1}{|J_t^1|} \sum_{j \in J_t^1} b_j(\vartheta, X_t) \quad \text{et} \quad c_J(Y_t) = \frac{1}{|J_t^1|} \sum_{j \in J_t^1} c_j(\vartheta, X_t), \quad (2.59)$$

où les égalités tiennent \mathbb{P} -p.s. Alors (2.58) peut être réécrite simplement comme

$$\dot{\mathcal{A}}_J(t) = -\mathbb{E}[A_t^J b_J(Y_t)] + \mathbb{E}[S_t^J a_J(Y_t)], \quad (2.60a)$$

$$\dot{\mathcal{R}}_J(t) = -\mathbb{E}[R_t^J c_J(Y_t)] + \mathbb{E}[A_t^J b_J(Y_t)], \quad (2.60b)$$

$$\dot{\mathcal{S}}_J(t) = -\mathbb{E}[S_t^J a_J(Y_t)] + \mathbb{E}[R_t^J c_J(Y_t)]. \quad (2.60c)$$

On obtient ainsi une description de la dynamique définie sur le graphe d'un point de vue purement macroscopique.

§ 2.5. Le problème de fermeture de moments

On a obtenu à la section précédente le système d'équations différentielles ordinaires (2.60) en supposant que pour chaque $x \in E$, les fonctions $a_j(\cdot, x)$, $b_j(\cdot, x)$ et $c_j(\cdot, x)$ sont intégrables sur l'espace des paramètres $(\Theta, \mathcal{G}, \mu)$, et en supposant que les taux de transition peuvent

être moyennés sur des populations du graphe de la façon décrite en (2.59). Le système (2.60) atteint alors les objectifs énoncés plus tôt : il fournit une description macroscopique moyenne de la dynamique sur le graphe qui ne dépend pas du choix de paramètres $\vartheta \in \Theta$. Cependant, un obstacle majeur reste à surmonter : ce système n'est pas autonome. En effet, les membres de droite des équations différentielles ne sont pas exprimés en fonction des espérances des composantes de l'état macroscopique Y_t .

Il est alors intéressant de remarquer que la démarche présentée aux deux sections précédentes peut être généralisée pour obtenir des équations différentielles qui décrivent les autres moments des composantes de l'état macroscopique. Par exemple, on peut s'intéresser à l'évolution des seconds moments, qu'on notera $E_{UV}^{JK}(t) := \mathbb{E}[U_t^J V_t^K]$ pour deux populations J et K du graphe, où U et V remplacent A , R ou S . On privilégiera cette notation lorsqu'on voudra voir les seconds moments comme des variables d'un système dynamique. En commençant par étudier comment évoluent les probabilités que deux noeuds soient dans deux états donnés, sous les mêmes suppositions que celles tenues pour obtenir le système (2.60), il est possible de voir que pour toutes populations J, K du graphe,

$$\dot{E}_{AA}^{JK}(t) = -\mathbb{E}[A_t^J A_t^K (b_J(Y_t) + b_K(Y_t))] + \mathbb{E}[A_t^J S_t^K a_K(Y_t)] + \mathbb{E}[S_t^J A_t^K a_J(Y_t)], \quad (2.61a)$$

$$\dot{E}_{RR}^{JK}(t) = -\mathbb{E}[R_t^J R_t^K (c_J(Y_t) + c_K(Y_t))] + \mathbb{E}[R_t^J A_t^K b_K(Y_t)] + \mathbb{E}[A_t^J R_t^K b_J(Y_t)], \quad (2.61b)$$

$$\dot{E}_{AR}^{JK}(t) = -\mathbb{E}[A_t^J R_t^K (b_J(Y_t) + c_K(Y_t))] + \mathbb{E}[A_t^J A_t^K b_K(Y_t)] + \mathbb{E}[S_t^J R_t^K a_J(Y_t)]. \quad (2.61c)$$

Les expressions analogues qui impliquent les fractions de populations dans l'état 0 suivent alors de (2.61) et du fait que $A_t^J + R_t^J + S_t^J \equiv 1$ pour tout $t \geq 0$ et pour toute population J . En effet, cette relation mène aux expressions

$$\dot{E}_{SS}^{JK}(t) = -\mathbb{E}[S_t^J S_t^K (a_J(Y_t) + a_K(Y_t))] + \mathbb{E}[S_t^J R_t^K c_K(Y_t)] + \mathbb{E}[R_t^J S_t^K c_J(Y_t)], \quad (2.61d)$$

$$\dot{E}_{AS}^{JK}(t) = -\mathbb{E}[A_t^J S_t^K (b_J(Y_t) + a_K(Y_t))] + \mathbb{E}[A_t^J R_t^K c_K(Y_t)] + \mathbb{E}[S_t^J S_t^K a_J(Y_t)], \quad (2.61e)$$

$$\dot{E}_{RS}^{JK}(t) = -\mathbb{E}[R_t^J S_t^K (c_J(Y_t) + a_K(Y_t))] + \mathbb{E}[R_t^J R_t^K c_K(Y_t)] + \mathbb{E}[A_t^J S_t^K b_J(Y_t)]. \quad (2.61f)$$

Les dérivées des seconds moments centrés des mêmes variables aléatoires, c'est-à-dire de leurs covariances, peuvent être décrites de la même façon. On notera ces covariances $C_{UV}^{JK}(t) := \text{Cov}[U_t^J, V_t^K]$ pour deux populations J et K du graphe, où U et V remplacent à nouveau A , R ou S . On peut effectivement tirer des systèmes (2.60) et (2.61) que

$$\begin{aligned} \dot{C}_{AA}^{JK}(t) &= -\text{Cov}[A_t^J, A_t^K b_K(Y_t)] - \text{Cov}[A_t^J b_J(Y_t), A_t^K] \\ &\quad + \text{Cov}[A_t^J, S_t^K a_K(Y_t)] + \text{Cov}[S_t^J a_J(Y_t), A_t^K], \end{aligned} \quad (2.62a)$$

$$\begin{aligned} \dot{C}_{AR}^{JK}(t) &= -\text{Cov}[A_t^J, R_t^K c_K(Y_t)] - \text{Cov}[A_t^J b_J(Y_t), R_t^K] \\ &\quad + \text{Cov}[A_t^J, A_t^K b_K(Y_t)] + \text{Cov}[S_t^J a_J(Y_t), R_t^K], \end{aligned} \quad (2.62b)$$

$$\begin{aligned} \dot{C}_{RR}^{JK}(t) &= -\text{Cov}[R_t^J, R_t^K c_K(Y_t)] - \text{Cov}[R_t^J c_J(Y_t), R_t^K] \\ &\quad + \text{Cov}[R_t^J, A_t^K b_K(Y_t)] + \text{Cov}[A_t^J b_J(Y_t), R_t^K]. \end{aligned} \quad (2.62c)$$

Des démarches complètes qui mènent aux résultats (2.61) et (2.62) sont données à l'annexe A. En principe, il serait possible d'obtenir des résultats semblables aussi pour les moments suivants de ces variables aléatoires, quoique le traitement serait de plus en plus fastidieux.

Comme on peut décrire les dérivées de tous les moments des composantes de l'état macroscopique, il est naturel de voir ceci comme un problème de fermeture de moments, qu'on a vu plus en détail à la section 1.3. Même dans le contexte du système (2.60)–(2.61) ou du système (2.60)–(2.62), il est difficile de proposer une solution générale à ce problème qui incluerait les moments d'ordre 2 et qui fonctionnerait peu importe la forme spécifique du système.

Cependant, il est toujours possible de considérer l'approximation de champ moyen du système, où tous les moments centrés d'ordre 2 et plus sont négligés. Avec cette supposition, on n'inclut aucun des systèmes (2.61) et (2.62), et on peut distribuer les espérances sur les différents facteurs des produits présents dans les équations (2.60). De plus, comme on néglige les variances de toutes les composantes de l'état macroscopique Y_t , alors on considère que $Y_t = \mathcal{Y}(t) := \mathbb{E}[Y_t]$ p.s., de sorte que par exemple

$$\mathbb{E}[A_t^J b_J(Y_t)] = \mathbb{E}[A_t^J] \mathbb{E}[b_J(Y_t)] = \mathcal{A}_J(t) b_J(\mathcal{Y}(t)). \quad (2.63)$$

En appliquant le même raisonnement aux autres produits, on obtient le système de champ moyen

$$\dot{\mathcal{A}}_J = -\mathcal{A}_J b_J(\mathcal{Y}) + \mathcal{S}_J a_J(\mathcal{Y}), \quad (2.64a)$$

$$\dot{\mathcal{R}}_J = -\mathcal{R}_J c_J(\mathcal{Y}) + \mathcal{A}_J b_J(\mathcal{Y}), \quad (2.64b)$$

$$\dot{\mathcal{S}}_J = -\mathcal{S}_J a_J(\mathcal{Y}) + \mathcal{R}_J c_J(\mathcal{Y}). \quad (2.64c)$$

Cette approche impose cependant de négliger l'essentiel des propriétés stochastiques de la dynamique initiale, puisque les variances des composantes de l'état macroscopique sont négligées.

Il est alors logique de chercher à aller plus loin, en incluant au moins les seconds moments des composantes de l'état macroscopique. L'objectif est alors de réécrire les membres de droite de (2.60) et de (2.61) en fonction des premiers et seconds moments des composantes de Y_t . Toutefois, on ne cherchera pas à obtenir un résultat général ici. En effet, les méthodes utilisées pour réaliser la fermeture de moments dépendent plutôt de la forme précise des taux de transition moyens a_J , b_J et c_J , et dépendent donc des applications.

Dans les chapitres suivants, ces idées seront développées dans deux applications particulières. On survolera d'abord les idées principales du problème dans le contexte d'un modèle épidémiologique au chapitre 3, puis on étudiera plus en détail le problème dans le contexte d'un modèle de dynamique neuronale au chapitre 4.

Chapitre 3

Application aux modèles épidémiologiques compartimentaux

Les modèles compartimentaux sont une famille de modèles épidémiologiques qui regroupe la plupart des modèles utilisés dans ce domaine [26, 46]. Les cas classiques ont été introduits en 1927 dans [57] par Kermack et McKendrick, qui ont poursuivi leur travaux au début des années 1930 dans [58, 59].

Dans ce chapitre, on s'intéresse particulièrement au modèle SIR, qui est le cas le plus classique. L'objectif de ce chapitre est de montrer que la dynamique du modèle SIR peut être vue comme un cas particulier de dynamique markovienne ternaire cyclique, et de mettre en évidence les défis que pose la modélisation d'un phénomène biologique par une telle dynamique. Le modèle SIR étant classique, le cas particulier du problème de fermeture de moments discuté à la section 2.5 pour ce modèle a déjà été étudié. On veut ainsi présenter quelques solutions de ce problème qui permettent d'inclure les moments d'ordre 2. Ce chapitre est entièrement indépendant du chapitre 4 et peut être sauté sans affecter la lecture. Ceci dit, l'étude du problème de fermeture de moments présentée dans le contexte du modèle SIR peut aider à saisir comment on aborde un tel problème de façon plus concrète.

Dans la première section, on donne une introduction plus détaillée des modèles compartimentaux en épidémiologie, en portant une attention particulière au modèle SIR. On montre ensuite dans la deuxième section qu'il est possible d'appliquer les résultats du chapitre 2 pour obtenir ce modèle. La troisième section est consacrée à une revue de quelques solutions possibles du problème de fermeture de moments pour le modèle SIR. Finalement, on conclut le chapitre par une brève discussion sur l'applicabilité des résultats du chapitre 2 à d'autres modèles compartimentaux que le modèle SIR.

§ 3.1. Les modèles compartimentaux

L'idée générale des modèles compartimentaux est de considérer une population d'individus à travers laquelle se propage une épidémie, et de la diviser en différents groupes, ou compartiments, dans lesquels les individus sont dans le même état relativement à la maladie

qui se propage [46]. Par exemple, on considère généralement un compartiment S pour des individus qui sont susceptibles d'être infectés par la maladie, et un compartiment I pour des individus infectés qui peuvent contaminer d'autres individus. On peut aussi considérer un compartiment E pour des individus exposés à la maladie mais qui ne sont pas encore contagieux, et un compartiment R pour des individus qui ont déjà été infectés et qui ne peuvent pas être réinfectés. Ce type de modèle est généralement nommé par un acronyme composé des lettres qui correspondent aux différents compartiments, en répétant la lettre S à la fin de l'acronyme lorsqu'un même individu peut être infecté plus d'une fois. On obtient ainsi les modèles SI, SIR, SEIR, SIS, SIRS et SEIRS, selon les compartiments choisis.

Il est aussi possible de considérer d'autres compartiments que les quatre mentionnés ici. Ces modèles peuvent également être adaptés pour tenir compte d'autres phénomènes, comme des naissances et des morts, de la vaccination, ou encore des variations saisonnières dans les contacts entre les individus [27].

§ 3.1.1. Le modèle SIR classique

Parmi les modèles compartimentaux, l'un des plus classiques est le modèle SIR présenté par Kermack et McKendrick dans [57] en 1927. Ce modèle permet de représenter la propagation d'une maladie infectieuse dans une population sous des hypothèses minimales. On présente dans cette section une brève dérivation du modèle, basée sur [70] et sur [46].

On considère une population de N individus. Pour modéliser une épidémie qui s'y propage, on la divise en trois compartiments. On considère d'abord un compartiment d'individus en santé, mais qui peuvent contracter la maladie. On dit généralement que ces individus sont *susceptibles*, et on note leur nombre au temps t par $S(t)$. Ensuite, on considère un compartiment d'individus qui sont *infectés* par la maladie et qui peuvent la transmettre aux individus susceptibles. On note par $I(t)$ le nombre d'individus infectés au temps t . Finalement, on considère un compartiment d'individus guéris, qui sont désormais immunisés et ne peuvent plus être infectés. On dit alors que ces individus sont *retirés*, et on note $R(t)$ leur nombre au temps t . On considère toujours que chaque individu est dans un et un seul de ces trois compartiments, de sorte que $S + I + R \equiv N$.

On veut ensuite décrire comment les trois quantités S , I et R évoluent dans le temps. Pour simplifier la situation, on suppose qu'elles varient de façon continue, même si dans la réalité elles variraient plutôt de façon discrète. L'idée générale de la modélisation, détaillée au prochain paragraphe, est résumée à la figure 3.1.

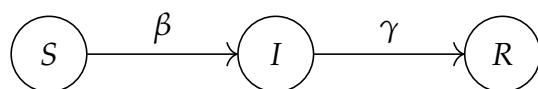


FIGURE 3.1 – Transitions entre les trois compartiments dans le modèle SIR.

D'abord, on suppose qu'un individu a β contacts suffisants avec d'autres individus par unité de temps, où un « contact suffisant » désigne un contact qui serait susceptible de causer la transmission de l'infection. En un temps t , un individu susceptible a donc $\beta I(t)/N$ contacts par unité de temps avec des individus infectés. Ainsi, au temps t , le nombre d'individus par unité de temps qui passent de l'état susceptible à l'état infecté est $\beta S(t)I(t)/N$. Comme ce phénomène est le seul qui fait varier le nombre d'individus susceptibles, on conclut que

$$\dot{S} = -\frac{\beta SI}{N}, \quad (3.1a)$$

où le point note la dérivée. On suppose ensuite que les individus infectés guérissent à un taux γ , c'est-à-dire que la maladie dure un temps de $1/\gamma$. Ainsi, au temps t , le nombre d'individus infectés qui guérissent par unité de temps est de $\gamma I(t)$. On trouve donc que

$$\dot{I} = \frac{\beta SI}{N} - \gamma I. \quad (3.1b)$$

Les individus infectés qui guérissent passent directement au compartiment des individus retirés, et ce phénomène est le seul qui fasse varier la quantité d'individus retirés. On déduit donc finalement que

$$\dot{R} = \gamma I. \quad (3.1c)$$

Le système (3.1) peut aussi être réécrit selon les fractions de la population totale qui sont dans chacun des trois compartiments. Pour ce faire, on définit

$$S(t) := \frac{S(t)}{N}, \quad I(t) := \frac{I(t)}{N} \quad \text{et} \quad R(t) := \frac{R(t)}{N}. \quad (3.2)$$

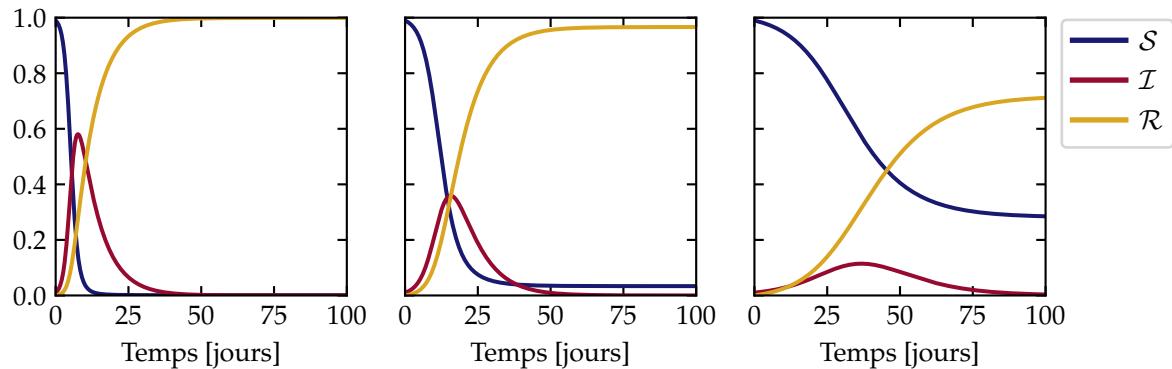


FIGURE 3.2 – Trois solutions du modèle SIR (3.3) pour $S(0) = 0,99$, $I(0) = 0,01$ et $R(0) = 0$. Dans les trois cas, le taux de guérison est $\gamma = 1/7$ jours $^{-1}$. La seule différence entre les trois solutions est la valeur du taux de transmission β : on a pris, dans l'ordre, $\beta = 1$, $\beta = 1/2$ et $\beta = 1/4$ jours $^{-1}$. Ceci donne une petite illustration de l'idée que de réduire les contacts (et donc de réduire la valeur de β) lors d'une épidémie est crucial afin d'aplatir la courbe de la proportion d'individus infectés.

En divisant par N les équations (3.1), on obtient alors le système

$$\dot{S} = -\beta S \mathcal{I}, \quad (3.3a)$$

$$\dot{\mathcal{I}} = \beta S \mathcal{I} - \gamma \mathcal{I}, \quad (3.3b)$$

$$\dot{\mathcal{R}} = \gamma \mathcal{I}. \quad (3.3c)$$

On donne des exemples de solutions de cette équation différentielle à la figure 3.2.

Le modèle SIR a été abondamment étudié au cours des dernières décennies, et ses propriétés mathématiques sont aujourd’hui bien connues. Une analyse élémentaire, mais très intéressante, est donnée au chapitre 2 de [70]. Plusieurs généralisations de ce modèle ont aussi été étudiées ; de bons exemples sont discutés dans [7, 46, 70].

§ 3.1.2. Le modèle SIRS

Le modèle SIRS est très semblable au modèle SIR, mais permet de modéliser une épidémie où l’immunité qui suit l’infection dure seulement un temps fini. Le passage des individus d’un compartiment à l’autre est alors représenté par le diagramme donné à la figure 3.3. Reprenant la modélisation présentée à la sous-section 3.1.1, on peut passer au modèle SIRS en supposant qu’un individu rétabli redevient susceptible au taux δ , c’est-à-dire que l’immunité dure un intervalle de temps $1/\delta$ plutôt que d’être infinie. Ceci a pour effet d’ajouter un terme $+\delta \mathcal{R}$ dans l’expression de la dérivée \dot{S} , et un terme $-\delta \mathcal{R}$ dans l’expression de la dérivée $\dot{\mathcal{R}}$. En modifiant ainsi le système (3.3), on obtient le système

$$\dot{S} = -\beta S \mathcal{I} + \delta \mathcal{R}, \quad (3.4a)$$

$$\dot{\mathcal{I}} = -\gamma \mathcal{I} + \beta S \mathcal{I}, \quad (3.4b)$$

$$\dot{\mathcal{R}} = -\delta \mathcal{R} + \gamma \mathcal{I}, \quad (3.4c)$$

qui décrit l’évolution des fractions d’individus qui se trouvent dans chacun des trois compartiments.

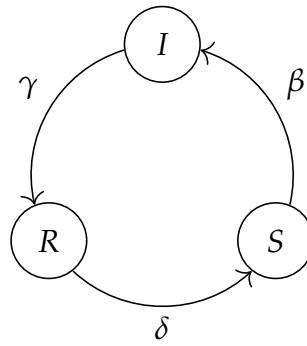


FIGURE 3.3 – Transitions entre les trois compartiments dans le modèle SIRS.

§ 3.2. Modélisation d’une épidémie par une dynamique MTC

La dérivation du modèle SIR présentée à la section 3.1.1 a l’avantage d’être claire et intuitive. Cependant, elle impose de considérer tous les individus de la population comme identiques, et de considérer que des contacts peuvent se produire de la même façon entre n’importe quelle paire d’individus. Afin d’alléger ces suppositions, on peut construire le modèle SIR à partir d’une description des interactions individuelles plutôt qu’à partir d’une

description globale de ce qui se produit dans la population. Par ailleurs, il est naturel de modéliser les interactions entre les individus et la transmission de l'infection comme des phénomènes aléatoires. Ceci mène à l'idée de modéliser une épidémie par un processus stochastique qui décrit les états de tous les individus de la population.

La formulation du modèle SIR à partir d'un processus stochastique est aujourd'hui une autre approche standard au modèle SIR, abordée par exemple dans [1, 54, 61, 60, 91, 94, 95]. On propose dans cette section une nouvelle construction du modèle SIR d'un point de vue stochastique, basée sur les résultats du chapitre 2. Même si cette construction du modèle est différente de ce qu'on retrouve dans la littérature puisqu'il s'agit de considérer un cas particulier d'une démarche plus générale, les résultats obtenus sont équivalents à ceux déjà connus.

§ 3.2.1. Modélisation de l'épidémie

On considère un ensemble de N individus à travers lequel se propage une épidémie. On vise à représenter la situation par une dynamique MTC sur un graphe aléatoire. On prend alors une matrice de poids W de dimension N et G le graphe aléatoire induit par cette matrice de poids. On interprète les nœuds de G comme les individus, et on interprète une composante W_{jk} de la matrice de poids comme le taux auquel les individus j et k ont des interactions qui mènent à l'infection de j par k , si k est infecté. Le rôle des poids est ainsi analogue au rôle du paramètre β dans le modèle SIR classique. On introduit également γ un paramètre d'ordre 1 sur G , qui joue le même rôle que dans le modèle SIR classique : la variable aléatoire γ_j décrit le taux auquel l'individu j guérit de l'infection lorsqu'il est infecté.

Il est simple d'expliciter un espace de probabilité sur lequel le graphe et ses paramètres peuvent être définis. En effet, soient $\mu_W: \mathcal{B}(\mathbb{R}^{N \times N}) \rightarrow [0, 1]$ la loi de probabilité voulue pour la matrice de poids et $\mu_\gamma: \mathcal{B}(\mathbb{R}^N) \rightarrow [0, 1]$ celle du paramètre γ . On peut alors prendre $\Theta := \mathbb{R}^{N^3}$, $\mathcal{G} := \mathcal{B}(\mathbb{R}^{N^3})$ et $\mu := \mu_\gamma \otimes \mu_W$, et l'espace de probabilité $(\Theta, \mathcal{G}, \mu)$ forme l'espace des paramètres du graphe. En voyant $\mathbb{R}^{N^3} = \mathbb{R}^N \times \mathbb{R}^{N \times N}$, les paramètres γ et W deviennent simplement des projections sur les deux composantes, qui sont évidemment \mathcal{G} -mesurables. On suppose également que les lois de probabilité des paramètres sont choisies de sorte qu'ils soient intégrables.

Afin de représenter les états des individus relativement à la maladie, on introduit un espace d'états ternaires (E, \mathcal{E}) pour G . On interprète alors :

- l'état 0 comme celui d'un individu *susceptible*,
- l'état 1 comme celui d'un individu *infecté*,
- l'état i comme celui d'un individu *retiré*.

On veut ensuite définir des taux de transition $a_j, b_j, c_j: \Theta \times E \rightarrow [0, \infty)$ entre ces trois états de façon à reproduire la dynamique schématisée à la figure 3.1.

D'abord, on sait qu'un individu susceptible j a, au taux W_{jk} , des contacts avec un autre individu k qui vont l'infecter si k est lui-même infecté. Ainsi, pour un état $x \in E$ du graphe, le taux a_j auquel j passe de l'état susceptible à l'état infecté est donné par

$$a_j(\vartheta, x) := \sum_{k=1}^N W_{jk}(\vartheta) \operatorname{Re} x_k. \quad (3.5)$$

Ensuite, on sait qu'indépendamment de l'état du graphe, l'individu j guérit au taux

$$b_j(\vartheta, x) := \gamma_j(\vartheta). \quad (3.6)$$

On suppose finalement qu'une fois qu'un individu a contracté la maladie et en est guéri, il devient immunisé. Ainsi, une fois un individu dans l'état retiré, il y reste. On prend donc simplement $c_j \equiv 0$ pour chaque nœud j du graphe.

Puisque γ et W sont \mathcal{G} -mesurables et intégrables, il est clair des expressions (3.5) et (3.6) que pour chaque $x \in E$, les projections $a_j(\cdot, x)$ et $b_j(\cdot, x)$ sont \mathcal{G} -mesurables et intégrables, de même que $c_j(\cdot, x)$ qui est identiquement nulle. Ainsi, par le théorème 2.3, si P est une mesure de probabilité sur (E, \mathcal{E}) , il existe pour chaque $\vartheta \in \Theta$ une mesure de probabilité \mathbb{P}^ϑ sur $(\Omega, \mathcal{F}) := (E, \mathcal{E})^{[0, \infty)}$ telle que les conditions *i*) à *iii*) du théorème sont respectées, et telle que le processus de projection de coordonnées $\{X_t\}_{t \geq 0}$ respecte les conditions (2.7) données au corollaire 1. De plus, par le corollaire 2, l'espace $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F})$ peut être muni d'une mesure de probabilité \mathbb{P} telle que pour toute variable aléatoire Z sur $\Theta \times \Omega$,

$$\mathbb{E}[Z] = \int_{\Theta} \int_{\Omega} Z(\vartheta, \omega) d\mathbb{P}^\vartheta(\omega) d\mu(\vartheta), \quad (3.7)$$

où \mathbb{E} est l'espérance par rapport à \mathbb{P} . Le processus $\{X_t\}_{t \geq 0}$ correspond alors à une dynamique MTC qui décrit l'état du graphe selon le temps.

§ 3.2.2. Passage à la description macroscopique de la dynamique

Afin d'obtenir une description macroscopique de la dynamique sur le graphe, qui modéliserait l'épidémie, on cherche à appliquer les résultats présentés à la section 2.4. Pour ce faire, on commence par diviser le graphe en une collection \mathcal{P} de n populations. Comme on a vu que les projections des taux de transition sur l'espace des paramètres sont intégrables, si l'on trouve des taux de transition moyens de la forme (2.59), alors la démarche décrite à la section 2.4 donnera directement une description macroscopique moyenne de la dynamique.

En reprenant la notation utilisée à la section 2.4, on cherche ainsi pour chaque population $J \in \mathcal{P}$ des fonctions $a_J, b_J, c_J : [0, 1]^{2n} \rightarrow [0, \infty)$ telles que pour tout $t \geq 0$,

$$a_J(Y_t) = \frac{1}{|J_t^0|} \sum_{j \in J_t^0} a_j(\vartheta, X_t) = \frac{1}{|J_t^0|} \sum_{j \in J_t^0} \sum_{k=1}^N W_{jk}(\vartheta) \operatorname{Re} X_t^k, \quad (3.8a)$$

$$b_J(Y_t) = \frac{1}{|J_t^1|} \sum_{j \in J_t^1} b_j(\vartheta, X_t) = \frac{1}{|J_t^1|} \sum_{j \in J_t^1} \gamma_j(\vartheta) \quad (3.8b)$$

et

$$c_J(Y_t) = \frac{1}{|J_t^1|} \sum_{j \in J_t^1} c_j(\vartheta, X_t) = 0, \quad (3.8c)$$

où les égalités tiennent \mathbb{P} -p.s. et où Y_t note l'état macroscopique qui correspond à X_t selon (2.49). Comme les c_j sont identiquement nuls, on peut prendre $c_J \equiv 0$ et la condition (3.8c) est directement respectée.

En général, il n'est pas facile d'assurer qu'il existe de telles fonctions a_J et b_J . Il est cependant possible de le faire si les paramètres sont déterminés sur \mathcal{P} (au sens de la définition 2.8). En effet, pour $J, K \in \mathcal{P}$, soient γ_J et W_{JK} les valeurs réelles telles que $\gamma_j = \gamma_J$ \mathbb{P} -p.s. et $W_{jk} = W_{JK}$ \mathbb{P} -p.s. pour $j \in J$ et $k \in K$. Alors

$$\frac{1}{|J_t^1|} \sum_{j \in J_t^1} \gamma_j = \frac{1}{|J_t^1|} \sum_{j \in J_t^1} \gamma_J = \gamma_J \quad \mathbb{P}\text{-p.s.} \quad (3.9)$$

et

$$\begin{aligned} \frac{1}{|J_t^0|} \sum_{j \in J_t^0} \sum_{k=1}^N W_{jk} \operatorname{Re} X_t^k &= \frac{1}{|J_t^0|} \sum_{j \in J_t^0} \sum_{K \in \mathcal{P}} \sum_{k \in K_t^1} W_{jk} \\ &= \frac{1}{|J_t^0|} \sum_{j \in J_t^0} \sum_{K \in \mathcal{P}} \sum_{k \in K_t^1} W_{JK} = \sum_{K \in \mathcal{P}} |K_t^1| W_{JK} = \sum_{K \in \mathcal{P}} |K| W_{JK} I_t^K \quad \mathbb{P}\text{-p.s.}, \end{aligned} \quad (3.10)$$

où I_t^K note la fraction de la population K qui se trouve dans l'état infecté (qui serait notée A_t^K à la section 2.4). Afin d'adopter une notation similaire à celle utilisée pour le modèle SIR classique, on introduit $\beta_{JK} := |K| W_{JK}$. Cette quantité est en effet analogue au paramètre β utilisé dans le modèle SIR classique, puisque W_{JK} représente le taux auquel un individu de la population J a des contacts avec un individu donné de la population K . Ainsi, β_{JK} représente le taux auquel un individu de la population J a des contacts avec les individus de la population K .

On peut ainsi prendre les taux moyens

$$a_J(Y_t) := \sum_{K \in \mathcal{P}} \beta_{JK} I_t^K \quad \text{et} \quad b_J(Y_t) := \gamma_J, \quad (3.11)$$

qui respectent les conditions (3.8). Avec ces taux, on obtient le système prévu par (2.60) à la section 2.4, qui prend ici la forme

$$\dot{S}_J(t) = -\mathbb{E} \left[S_t^J \sum_{K \in \mathcal{P}} \beta_{JK} I_t^K \right] = - \sum_{K \in \mathcal{P}} \beta_{JK} \mathbb{E}[S_t^J I_t^K], \quad (3.12a)$$

$$\dot{I}_J(t) = -\mathbb{E}[I_t^J \gamma_J] + \mathbb{E} \left[S_t^J \sum_{K \in \mathcal{P}} \beta_{JK} I_t^K \right] = -\gamma_J \mathcal{I}_J(t) + \sum_{K \in \mathcal{P}} \beta_{JK} \mathbb{E}[S_t^J I_t^K], \quad (3.12b)$$

$$\dot{\mathcal{R}}_J(t) = \mathbb{E}[I_t^J \gamma_J] = \gamma_J \mathcal{I}_J(t), \quad (3.12c)$$

où l'on a remplacé la lettre A par I dans les notations pour être plus cohérent avec le modèle SIR classique.

On peut ensuite obtenir l'approximation de champ moyen de ce système. Il est intéressant de remarquer qu'ici, on n'a pas à négliger les variances des fractions de populations pour l'obtenir. En effet, il suffit de négliger les corrélations entre des fractions susceptibles et infectées des populations. Dans ce cas, on considère que $\mathbb{E}[S_t^J I_t^K] = S_J(t) I_K(t)$ pour tout $t \geq 0$ et pour n'importe quelle paire de populations $J, K \in \mathcal{P}$. Le système (3.12) est alors simplifié comme

$$\dot{S}_J = - \sum_{K \in \mathcal{P}} \beta_{JK} S_J I_K, \quad (3.13a)$$

$$\dot{I}_J = -\gamma_J I_J + \sum_{K \in \mathcal{P}} \beta_{JK} S_J I_K, \quad (3.13b)$$

$$\dot{R}_J = \gamma_J I_J, \quad (3.13c)$$

qui donne une équation différentielle autonome. Il est alors intéressant de remarquer que, si l'on ne considère qu'une seule population, ce système est réduit à

$$\dot{S} = -\beta S I, \quad (3.14a)$$

$$\dot{I} = -\gamma I + \beta S I, \quad (3.14b)$$

$$\dot{R} = \gamma I, \quad (3.14c)$$

où l'on a omis les indices qui réfèrent à l'unique population. On retrouve donc le système (3.3), qui correspond au modèle SIR classique ! Le système (3.13) généralise donc le modèle SIR en permettant différentes populations qui auraient des contacts plus ou moins forts.

§ 3.2.3. Remarques sur les suppositions qui mènent au système (3.12)

Pour obtenir le système (3.12), on a supposé que les paramètres étaient déterminés sur \mathcal{P} . Comme cette supposition est assez lourde, il semble intéressant de tenter de l'alléger.

On suppose donc que les paramètres sont plutôt homogènes sur \mathcal{P} (au sens de la définition 2.7). Si le graphe est divisé en un petit nombre de très grandes populations, alors

$$\frac{1}{|J_t^1|} \sum_{j \in J_t^1} \gamma_j \quad \text{et} \quad \frac{1}{|J_t^0|} \sum_{j \in J_t^0} \sum_{k=1}^N W_{jk} \operatorname{Re} X_t^k = \frac{1}{|J_t^0|} \sum_{j \in J_t^0} \sum_{K \in \mathcal{P}} \sum_{k \in K_t^1} W_{jk} \quad (3.15)$$

représentent les moyennes arithmétiques de variables aléatoires indépendantes et identiquement distribuées sur des ensembles qui sont potentiellement très grands. Il semblerait alors approprié de tenter de considérer la limite où la taille de chaque population tend vers l'infini, afin de remplacer (3.9) et (3.10) par des utilisations de la loi des grands nombres.

Cependant, il est difficile de considérer correctement une telle limite. En effet, on a pu construire un espace de probabilité $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F}, P)$ à la section 2.2 en fixant un nombre de nœuds N pour le graphe. Pour différentes tailles de populations, on a donc différents espaces de probabilité, et il n'est pas clair de quelle façon on pourrait les relier. De plus,

les ensembles J_t^0 et J_t^1 sur lesquels les taux de transition sont sommés dans (3.15) sont eux-mêmes aléatoires. Pour pouvoir utiliser une forme de loi des grands nombres, il faudrait donc arriver à formuler une limite où les tailles de J_t^0 et de J_t^1 deviennent arbitrairement grandes avec une probabilité élevée.

Ainsi, il ne serait certainement pas simple d'appliquer une forme de loi des grands nombres pour obtenir des analogues de (3.9) et de (3.10) dans le cas où les paramètres sont homogènes sur \mathcal{P} . Néanmoins, si une population J est très grande, il semble raisonnable de supposer que si $\mathcal{I}_J(t) \neq 0$, alors l'ensemble J_t^1 est aussi très grand avec une probabilité élevée. Par contre, si $\mathcal{I}_J(t) = 0$, alors l'ensemble J_t^1 est vide presque sûrement. Il semble donc raisonnable d'approximer que

$$\frac{1}{|J_t^1|} \sum_{j \in J_t^1} \gamma_j \approx \gamma_J := \mathbb{E}[\gamma_j] \text{ pour } j \in J, \quad (3.16)$$

du moins dans le contexte où l'on veut motiver (3.12) pour des paramètres homogènes sur \mathcal{P} . De la même façon, il semble raisonnable d'approximer que

$$\frac{1}{|J_t^0|} \sum_{j \in J_t^0} \sum_{k=1}^N W_{jk}(\vartheta) \operatorname{Re} X_t^k \approx \sum_{K \in \mathcal{P}} |K| W_{JK} I_t^K(t) = \sum_{K \in \mathcal{P}} \beta_{JK} I_t^K(t), \quad (3.17)$$

où ici $W_{JK} := \mathbb{E}[W_{jk}]$ pour $j \in J$ et $k \in K$, et où l'on prend toujours $\beta_{JK} = |K| W_{JK}$.

Dans le cas de paramètres homogènes sur \mathcal{P} , on verra donc le système (3.12) comme un modèle macroscopique approximatif.

§ 3.3. Fermetures de moments au-delà de l'approximation de champ moyen

Afin de décrire la propagation d'une épidémie à la manière du modèle SIR, on a obtenu à la section précédente le système (3.12), donné par

$$\dot{S}_J = - \sum_{K \in \mathcal{P}} \beta_{JK} E_{SI}^{JK}, \quad (3.12a)$$

$$\dot{I}_J = -\gamma_J \mathcal{I}_J + \sum_{K \in \mathcal{P}} \beta_{JK} E_{SI}^{JK}, \quad (3.12b)$$

$$\dot{\mathcal{R}}_J = \gamma_J \mathcal{I}_J, \quad (3.12c)$$

où l'on reprend la notation $E_{UV}^{JK}(t) := \mathbb{E}[U_t^J V_t^K]$ utilisée à la section 2.5 pour désigner les seconds moments. Pour considérer davantage l'aspect stochastique de la dynamique que dans l'approximation de champ moyen (3.13), on peut tenter de considérer aussi les seconds moments plutôt que de les approximer à partir des premiers. Pour ce faire, il faut d'abord compléter le système d'équations différentielles précédent avec la dérivée des seconds moments. Comme $S_t^J + I_t^J + R_t^J \equiv 1$ pour tout $t \geq 0$ et pour toute population J , les espérances

de produits de S_t^J et de I_t^J suffisent à déterminer tous les seconds moments. Avec les taux de transition moyens donnés en (3.11), les expressions (2.61e), (2.61d) et (2.61a) donnent

$$\dot{E}_{SI}^{JK}(t) = -\gamma_K E_{SI}^{JK}(t) - \sum_{L \in \mathcal{P}} \beta_{JL} \mathbb{E}[S_t^J I_t^K I_t^L] + \sum_{L \in \mathcal{P}} \beta_{KL} \mathbb{E}[S_t^J S_t^K I_t^L], \quad (3.18a)$$

$$\dot{E}_{SS}^{JK}(t) = - \sum_{L \in \mathcal{P}} (\beta_{JL} + \beta_{KL}) \mathbb{E}[S_t^J S_t^K I_t^L], \quad (3.18b)$$

$$\dot{E}_{II}^{JK}(t) = -(\gamma_J + \gamma_K) E_{II}^{JK}(t) + \sum_{L \in \mathcal{P}} \beta_{KL} \mathbb{E}[S_t^K I_t^J I_t^L] + \sum_{L \in \mathcal{P}} \beta_{JL} \mathbb{E}[S_t^J I_t^K I_t^L] \quad (3.18c)$$

où, comme dans la section précédente, on a remplacé la lettre A par la lettre I par cohérence avec le modèle SIR classique. On veut ensuite imposer des approximations supplémentaires afin d'obtenir une équation différentielle autonome à partir du système (3.12)–(3.18). Pour ce faire, il suffit d'approximer les moments d'ordre 3 en termes des moments d'ordre 2. Ce problème de fermeture de moments, de même que ses analogues pour d'autres modèles épidémiologiques compartimentaux, est très étudié depuis les années 1990. Plusieurs solutions sont discutées dans [60, 78, 79]. Dans le reste de la section, on donne deux exemples importants de solutions à ce problème.

§ 3.3.1. L'approximation des paires

Une solution populaire à ce problème de fermeture de moments est l'approximation des paires¹, introduite par Matsuda *et al.* en 1992 dans [71]. Cette méthode a été proposée à l'origine pour étudier des dynamiques de populations de type Lotka–Volterra en écologie², mais avec l'idée qu'elle puisse s'appliquer également à des modèles épidémiologiques compartimentaux. Elle a effectivement été reprise dans les deux disciplines, par exemple dans [54, 56, 55, 61, 67, 83, 88, 91]. Cette méthode est aujourd'hui standard, et discutée jusque dans des manuels de référence plus généraux comme [60, 82].

L'idée de cette approximation dans [71] provient de probabilités de trois événements simultanés qui peuvent être vues comme des espérances de produits de trois variables aléatoires binaires. On considère j , k et ℓ , trois noeuds distincts du graphe, et on suppose qu'il existe des liens entre j et k et entre j et ℓ . On cherche alors à approximer l'espérance $\mathbb{E}[(1 - |X_t^j|) \text{Re}(X_t^k) \text{Re}(X_t^\ell)]$. Puisque cette espérance est celle du produit de trois variables aléatoires binaires qui valent 0 ou 1, elle peut directement être réécrite comme une probabilité. En effet,

$$\mathbb{E}[(1 - |X_t^j|) \text{Re}(X_t^k) \text{Re}(X_t^\ell)] = \mathbb{P}[X_t^j = 0, X_t^k = 1, X_t^\ell = 1] \quad (3.19)$$

$$= \mathbb{P}[X_t^\ell = 1 | X_t^j = 0, X_t^k = 1] \mathbb{P}[X_t^j = 0, X_t^k = 1]. \quad (3.20)$$

1. Traduction libre de l'expression « *pair approximation* », souvent utilisée dans la littérature anglophone.

2. Voir l'exemple 1.1 pour la version classique de ce modèle.

Puisqu'on suppose qu'il existe un lien entre ℓ et j mais pas entre ℓ et k , on peut penser que l'état de j a davantage d'influence sur l'état de ℓ que l'état de k en a. Cette intuition justifie d'approximer que $\mathbb{P}[X_t^\ell = 1 | X_t^j = 0, X_t^k = 1] \approx \mathbb{P}[X_t^\ell = 1 | X_t^j = 0]$. Avec cette approximation, la relation précédente devient

$$\mathbb{E}[(1 - |X_t^j|) \operatorname{Re}(X_t^k) \operatorname{Re}(X_t^\ell)] \approx \mathbb{P}[X_t^\ell = 1 | X_t^k = 0] \mathbb{P}[X_t^j = 0, X_t^k = 1] \quad (3.21)$$

$$= \frac{\mathbb{P}[X_t^\ell = 1, X_t^k = 0]}{\mathbb{P}[X_t^k = 0]} \mathbb{P}[X_t^j = 0, X_t^k = 1] \quad (3.22)$$

$$= \frac{\mathbb{E}[(1 - |X_t^j|) \operatorname{Re}(X_t^\ell)] \mathbb{E}[(1 - |X_t^j|) \operatorname{Re}(X_t^k)]}{\mathbb{E}[1 - |X_t^k|]}. \quad (3.23)$$

On a ainsi approximé un moment d'ordre 3 en fonction de moments d'ordre 1 et 2 seulement.

On peut utiliser le même genre d'intuition pour justifier l'approximation lorsque les variables aléatoires d'intérêt ne sont pas binaires. Dans ce cas, la justification peut être basée sur la supposition que des trois interactions par paires entre les variables aléatoires du produit, deux soient plus importantes que la troisième. On considère d'abord le cas des espérances de la forme $\mathbb{E}[S_t^J I_t^K I_t^L]$ qui surviennent dans l'équation (3.18a). On sait que

$$\mathbb{E}[S_t^J I_t^K I_t^L] = \frac{1}{|J||K||L|} \sum_{j \in J} \sum_{k \in K} \sum_{\ell \in L} \mathbb{E}[(1 - |X_t^j|) \operatorname{Re}(X_t^k) \operatorname{Re}(X_t^\ell)] \quad (3.24)$$

$$= \frac{1}{|J||K||L|} \sum_{j \in J} \sum_{k \in K} \sum_{\ell \in L} \mathbb{P}[X_t^j = 0, X_t^k = 1, X_t^\ell = 1], \quad (3.25)$$

c'est-à-dire que $\mathbb{E}[S_t^J I_t^K I_t^L]$ donne la probabilité que si l'on tire trois individus j , k et ℓ respectivement dans les populations J , K et L , j soit susceptible alors que k et ℓ soient infectés. L'espérance d'intérêt survient dans l'expression de la dérivée de $\mathbb{E}[S_t^J I_t^K]$, spécifiquement au terme qui tient compte de l'infection d'individus de la population J par des individus de la population L . Il est donc logique de penser que le lien entre les populations J et K et celui entre les populations J et L priment sur le lien entre les populations K et L . Ceci motive l'idée de séparer l'espérance du produit sous la forme

$$\mathbb{E}[S_t^J I_t^K I_t^L] \approx \frac{\left(\begin{array}{c} \text{Probabilité que si l'on tire } j \text{ dans } J \\ \text{et } k \text{ dans } K, X_t^j = 0 \text{ et } X_t^k = 1 \end{array} \right) \left(\begin{array}{c} \text{Probabilité que si l'on tire } j \text{ dans } J \\ \text{et } \ell \text{ dans } L, X_t^j = 0 \text{ et } X_t^\ell = 1 \end{array} \right)}{\text{Probabilité que si l'on tire } j \text{ dans } J, X_t^j = 0}, \quad (3.26)$$

où l'on divise les deux probabilités au numérateur par la probabilité de tirer un individu susceptible dans J , puisque celle-ci est comptée deux fois au numérateur. Ces trois probabilités se réécrivent comme des espérances des fractions de populations en suivant le même principe qu'en (3.25). En effet, elles se traduisent de sorte que

$$\mathbb{E}[S_t^J I_t^K I_t^L] \approx \frac{\mathbb{E}[S_t^J I_t^K] \mathbb{E}[S_t^J I_t^L]}{\mathbb{E}[S_t^J]}. \quad (3.27)$$

On obtient alors une approximation d'un moment d'ordre 3 en termes de moments d'ordre 1 et 2, de la même façon qu'en (3.23). Certains auteurs, comme [55, 56, 83], proposent des corrections supplémentaires à cette approximation afin de mieux tenir compte de la structure du graphe, mais on n'ira pas plus loin ici.

Le même raisonnement peut être appliqué aux autres espérances de produits de trois variables aléatoires dans (3.18a), qui sont de la forme $\mathbb{E}[S_t^J S_t^K I_t^L]$. Cette espérance survient à nouveau dans l'expression de la dérivée de $\mathbb{E}[S_t^J I_t^K]$, cette fois-ci au terme qui tient compte de l'infection d'individus de la population K par des individus de la population L . On suppose donc que dans ce contexte, le lien entre les populations J et L est de moindre importance que les liens entre J et K et entre K et L , et on approxime que

$$\mathbb{E}[S_t^J S_t^K I_t^L] \approx \frac{\mathbb{E}[S_t^J S_t^K] \mathbb{E}[S_t^K I_t^L]}{\mathbb{E}[S_t^K]}. \quad (3.28)$$

Les approximations (3.27) et (3.28) permettent d'exprimer les dérivées des espérances de la forme $\mathbb{E}[S_t^J I_t^K]$ en termes de moments d'ordre 1 et 2 seulement, mais elles font également intervenir des espérances de la forme $\mathbb{E}[S_t^J S_t^K]$. Pour obtenir une équation différentielle autonome, on doit donc aussi approximer leurs dérivées, qui sont données par (3.18b). Ces dérivées contiennent seulement des espérances de la forme $\mathbb{E}[S_t^J S_t^K I_t^L]$, mais dans deux contextes distincts : l'infection d'individus de J par des individus de L , et l'infection d'individus de K par des individus de L . Suivant la même logique que plus tôt, dans ces deux contextes on approxime respectivement que

$$\mathbb{E}[S_t^J S_t^K I_t^L] \approx \frac{\mathbb{E}[S_t^J S_t^K] \mathbb{E}[S_t^J I_t^L]}{\mathbb{E}[S_t^J]} \quad \text{et que} \quad \mathbb{E}[S_t^J S_t^K I_t^L] \approx \frac{\mathbb{E}[S_t^J S_t^K] \mathbb{E}[S_t^K I_t^L]}{\mathbb{E}[S_t^K]}. \quad (3.29)$$

Puisque les approximations décrites plus haut ne font intervenir que des seconds moments de forme $\mathbb{E}[S_t^J I_t^K]$ ou $\mathbb{E}[S_t^J S_t^K]$, ceux-ci suffisent à rendre autonome le système (3.12). En utilisant les approximations (3.27) et (3.28) dans (3.18a) et en utilisant l'approximation (3.29) dans (3.18b), le système (3.12)–(3.18) devient autonome. Reprenant les notations utilisées plus tôt pour les espérances des fractions de populations et de leurs produits, on trouve alors que

$$\dot{\mathcal{S}}_J = - \sum_{L \in \mathcal{P}} \beta_{JL} E_{SI}^{JL}, \quad (3.30a)$$

$$\dot{\mathcal{I}}_J = -\gamma_J \mathcal{I}_J + \sum_{L \in \mathcal{P}} \beta_{JL} E_{SI}^{JL}, \quad (3.30b)$$

$$\dot{\mathcal{R}}_J = \gamma_J \mathcal{I}_J, \quad (3.30c)$$

$$\dot{E}_{SI}^{JK} = -\gamma_K E_{SI}^{JK} - \sum_{L \in \mathcal{P}} \beta_{JL} \frac{E_{SI}^{JK} E_{SI}^{JL}}{\mathcal{S}_J} + \sum_{L \in \mathcal{P}} \beta_{KL} \frac{E_{SS}^{JK} E_{SI}^{KL}}{\mathcal{S}_K}, \quad (3.30d)$$

$$\dot{E}_{SS}^{JK} = - \sum_{L \in \mathcal{P}} \beta_{JL} \frac{E_{SS}^{JK} E_{SI}^{JL}}{\mathcal{S}_J} - \sum_{L \in \mathcal{P}} \beta_{KL} \frac{E_{SS}^{JK} E_{SI}^{KL}}{\mathcal{S}_K}. \quad (3.30e)$$

Dans le cas où le graphe n'est composé que d'une population, ce système est réduit à

$$\dot{\mathcal{S}} = -\beta E_{SI}, \quad (3.31a)$$

$$\dot{\mathcal{I}} = -\gamma \mathcal{I} + \beta E_{SI}, \quad (3.31b)$$

$$\dot{\mathcal{R}} = \gamma \mathcal{I}, \quad (3.31c)$$

$$\dot{E}_{SI} = -\gamma E_{SI} + \beta \frac{E_{SI}}{\mathcal{S}} (E_{SS} - E_{SI}), \quad (3.31d)$$

$$\dot{E}_{SS} = -2\beta \frac{E_{SS} E_{SI}}{\mathcal{S}}. \quad (3.31e)$$

Les propriétés de ce système ont déjà été largement étudiées, par exemple dans [55, 83] ou plus récemment dans [105].

§ 3.3.2. L'approximation de champ moyen de deuxième ordre

Une autre solution à ce problème de fermeture de moments peut être obtenue à partir de l'approximation proposée par Cator et Van Mieghem dans [16] dans le contexte du modèle SIS. L'idée de l'approximation correspond simplement à réécrire l'espérance d'un produit de trois variables aléatoires comme une probabilité d'observer trois événements simultanément, puis séparer cette probabilité en deux à partir d'un conditionnement, qui est finalement négligé.

Il est simple d'appliquer cette idée afin d'approximer les espérances de la forme $\mathbb{E}[S_t^J I_t^K I_t^L]$ qui surviennent dans l'équation (3.18a). Pour simplifier, on suppose que les trois populations J , K et L sont distinctes. Comme plus tôt, on sait que

$$\mathbb{E}[S_t^J I_t^K I_t^L] = \frac{1}{|J||K||L|} \sum_{j \in J} \sum_{k \in K} \sum_{\ell \in L} \mathbb{E}[(1 - |X_t^j|) \operatorname{Re}(X_t^k) \operatorname{Re}(X_t^\ell)] \quad (3.32)$$

$$= \frac{1}{|J||K||L|} \sum_{j \in J} \sum_{k \in K} \sum_{\ell \in L} \mathbb{P}[X_t^j = 0, X_t^k = 1, X_t^\ell = 1]. \quad (3.33)$$

Ces probabilités peuvent alors être séparées en produits, pour donner que

$$\mathbb{E}[S_t^J I_t^K I_t^L] = \frac{1}{|J||K||L|} \sum_{j \in J} \sum_{k \in K} \sum_{\ell \in L} \mathbb{P}[X_t^\ell = 1 | X_t^j = 0, X_t^k = 1] \mathbb{P}[X_t^j = 0, X_t^k = 1]. \quad (3.34)$$

C'est ici qu'on est en mesure d'utiliser l'idée de [16], qui est de négliger le conditionnement sur les états de j et de k . On obtient alors que

$$\mathbb{E}[S_t^J I_t^K I_t^L] \approx \frac{1}{|J||K||L|} \sum_{j \in J} \sum_{k \in K} \sum_{\ell \in L} \mathbb{P}[X_t^\ell = 1] \mathbb{P}[X_t^j = 0, X_t^k = 1] \quad (3.35)$$

$$= \left(\frac{1}{|L|} \sum_{\ell \in L} \mathbb{E}[\operatorname{Re} X_t^\ell] \right) \left(\frac{1}{|J||K|} \sum_{j \in J} \sum_{k \in K} \mathbb{E}[(1 - |X_t^j|) \operatorname{Re} X_t^k] \right) \quad (3.36)$$

$$= \mathbb{E}[I_t^L] \mathbb{E}[S_t^J I_t^K]. \quad (3.37)$$

Dans le cas où les populations J , K et L ne seraient pas toutes distinctes, on pourrait ajouter des corrections à cette approximation. Par exemple, si $K = L$, alors dans (3.33) il se peut que $k = \ell$. Dans ce cas, les événements $X_t^k = 1$ et $X_t^\ell = 1$ sont en fait un seul événement, et le conditionnement en (3.34) n'a pas d'effet : la probabilité $\mathbb{P}[X_t^\ell = 1 | X_t^j = 0, X_t^k = 1]$ vaut directement 1. Il faudrait donc séparer la somme en deux pour traiter séparément les termes où $k = \ell$, et ajouter une correction en (3.35) afin de pouvoir retrouver les fractions de populations en (3.37). Cette correction serait toutefois proportionnelle à $1/|K|$, puisqu'il y a $|K|$ paires contenant deux fois le même nœud parmi les $|K|^2$ paires $(k, \ell) \in K^2$. En considérant de grandes populations, on choisit donc ici de négliger de telles corrections et d'utiliser l'approximation (3.37) même lorsque les populations ne sont pas distinctes.

En suivant le même principe, on peut approximer que

$$\mathbb{E}[S_t^J S_t^K I_t^L] \approx \mathbb{E}[I_t^L] \mathbb{E}[S_t^J S_t^K]. \quad (3.38)$$

Avec ces approximations et avec la notation utilisée plus tôt, le système (3.12)–(3.18) devient

$$\dot{S}_J = - \sum_{L \in \mathcal{P}} \beta_{JL} E_{SI}^{JL}, \quad (3.39a)$$

$$\dot{I}_J = -\gamma_J I_J + \sum_{L \in \mathcal{P}} \beta_{JL} E_{SI}^{JL}, \quad (3.39b)$$

$$\dot{\mathcal{R}}_J = \gamma_J I_J, \quad (3.39c)$$

$$\dot{E}_{SI}^{JK} = -\gamma_K E_{SI}^{JK} - \sum_{L \in \mathcal{P}} \beta_{JL} E_{SI}^{JK} I_L + \sum_{L \in \mathcal{P}} \beta_{KL} E_{SS}^{JK} I_L, \quad (3.39d)$$

$$\dot{E}_{SS}^{JK} = - \sum_{L \in \mathcal{P}} (\beta_{JL} + \beta_{KL}) E_{SS}^{JK} I_L, \quad (3.39e)$$

qui donne bien une équation différentielle autonome.

Ces approximations sont étudiées par Cator et Van Mieghem dans [16] pour le modèle SIS étudié du point de vue microscopique, où les variables aléatoires d'intérêt représentent les états des individus et non des populations. Dans ce contexte, ils montrent que le système résultant de ces approximations permet d'améliorer les prédictions du modèle de champ moyen lorsque le nombre N d'individus est assez grand. Toutefois, ils soutiennent aussi que lorsque N est très petit (inférieur ou égal à 5), il arrive que les solutions de leur système ne soient pas vraisemblables, car certaines composantes se mettent à grandir sans limites alors qu'elles devraient rester bornées. Ils argumentent alors que cette situation est causée par le fait qu'étant donnés les premiers et les seconds moments, il est possible qu'il n'existe pas de loi de probabilité qui respecte la condition imposée par l'approximation. Il est intéressant de remarquer que ceci ne peut pas survenir dans un modèle de champ moyen, puisque pour n'importe quelles valeurs choisies pour les premiers moments, il existe certainement une loi de probabilité où tous les moments centrés sont nuls.

§ 3.3.3. Bref survol d'autres méthodes

Il y aurait encore d'autres façons de trouver une solution au problème de fermeture de moments pour le modèle SIR. Par exemple, certains auteurs comme [55, 91, 90] utilisent une approche semblable à l'approximation des paires, mais sans supposer que le lien entre l'une des paires est moins fort que les autres. Avec cette idée, l'approximation (3.27) est remplacée par

$$\mathbb{E}[S_t^J I_t^K I_t^L] \approx \frac{\mathbb{E}[S_t^J I_t^K] \mathbb{E}[S_t^J I_t^L] \mathbb{E}[I_t^K I_t^L]}{\mathbb{E}[S_t^J] \mathbb{E}[I_t^K] \mathbb{E}[I_t^L]}. \quad (3.40)$$

Les autres espérances de produits de trois facteurs sont alors approximés de façon semblable, ce qui mène à un système un peu différent de (3.30).

Comme on peut le voir dans cette section, les méthodes utilisées pour solutionner ce problème de fermeture de moments sont généralement basées sur des arguments plus intuitifs que mathématiques. Certains auteurs proposent des méthodes plus rigoureuses. Par exemple, [85] propose une méthode basée sur une maximisation d'entropie. Quoique cette méthode soit plus attrayante d'un point de vue théorique puisqu'elle n'implique pas d'ajouter des suppositions pour approximer les troisièmes moments, elle est en général plus difficile à appliquer [79].

Finalement, il est important de remarquer qu'il serait aussi possible de construire un modèle plus complet encore en introduisant non seulement les moments d'ordre 2, mais aussi ceux d'ordre supérieur. House *et al.* ont présenté dans [52] une version du modèle SIS qui inclut jusqu'aux moments d'ordre 3, et leurs méthodes pourraient être appliquées au modèle SIR également.

§ 3.4. Applicabilité à d'autres modèles compartimentaux

Les idées présentées aux sections 3.2 et 3.3 ont été présentées seulement dans le contexte du modèle SIR, mais elles pourraient aussi être appliquées à d'autres modèles compartimentaux en épidémiologie, particulièrement au modèle SIRS.

La modélisation d'une épidémie par une dynamique MTC proposée à la section 3.2 peut facilement être adaptée de façon à reproduire le modèle SIRS plutôt que le modèle SIR. Pour ce faire, on introduit sur le graphe un autre paramètre d'ordre 1, noté δ , en supposant que sa loi de probabilité est choisie de sorte qu'il soit intégrable. Il est alors plus simple d'adapter la définition de l'espace des paramètres en suivant la même logique que celle décrite à la section 3.2.1. On définit ensuite les taux de transition c_j comme $c_j(\vartheta, x) := \delta_j(\vartheta)$ plutôt que de les définir identiquement nuls. Comme pour les autres paramètres, si l'on suppose que le paramètre δ est déterminé ou homogène sur la division \mathcal{P} du graphe en populations, on peut prendre les taux de transition moyens $c_J \equiv \delta_J := \mathbb{E}[\delta_j]$ pour $j \in J \in \mathcal{P}$.

Le système prévu par (2.60) à la section 2.4 prend alors la forme

$$\dot{S}_J(t) = - \sum_{K \in \mathcal{P}} \beta_{JK} \mathbb{E}[S_t^J I_t^K], \quad (3.41a)$$

$$\dot{I}_J(t) = -\gamma_J \mathcal{I}_J(t) + \sum_{K \in \mathcal{P}} \beta_{JK} \mathbb{E}[S_t^J I_t^K], \quad (3.41b)$$

$$\dot{\mathcal{R}}_J(t) = -\delta_J \mathcal{R}_J + \gamma_J \mathcal{I}_J(t). \quad (3.41c)$$

Dans le cas d'une seule population et dans l'approximation de champ moyen, ceci devient le système (3.4) du modèle SIRS classique. Les méthodes discutées à la section 3.3 pourraient également être utilisées pour rendre ce système autonome sans négliger les variances.

Finalement, il est intéressant de remarquer que les autres modèles compartimentaux utilisés en épidémiologie pourraient tous être étudiés par des méthodes similaires à celles présentées dans ce chapitre. Il faudrait toutefois modifier l'espace des états possibles de façon à ce que chaque nœud du graphe puisse prendre un nombre d'états différent de trois.

Chapitre 4

Application aux réseaux de neurones biologiques

Plusieurs modèles de champ moyen ont été proposés afin de décrire la dynamique d'un réseau de neurones biologiques. Un exemple classique très important est le modèle de Wilson–Cowan [101]. Une limite importante de ce modèle est qu'il néglige les corrélations qui pourraient exister entre les états des neurones, et beaucoup d'efforts sont mis en recherche afin de permettre de telles corrélations [10, 17]. L'objectif principal de ce chapitre est de proposer de nouvelles méthodes qui permettraient d'inclure de telles corrélations. Pour ce faire, on propose de modéliser le comportement d'un réseau de neurones biologiques par une dynamique markovienne ternaire cyclique, à partir des résultats du chapitre 2.

Dans la première section, on présente une brève dérivation du modèle de Wilson–Cowan, ainsi qu'un bref historique des travaux qui y ont mené et une discussion sur les liens qu'il a avec d'autres modèles. Dans la deuxième section, on applique les résultats du chapitre 2 afin de modéliser la dynamique d'un réseau de neurones biologiques. On obtiendra alors des équations différentielles qui posent un problème de fermeture de moments qui, à la connaissance de l'auteur de ce mémoire, n'a jamais été étudié auparavant. On verra cependant que l'approximation de champ moyen obtenue de ce problème n'est qu'une légère généralisation du modèle de Wilson–Cowan. Ces deux modèles sont alors comparés à la troisième section. On y étudie notamment deux exemples, illustrés respectivement aux figures 4.3 et 4.7a, où les deux modèles prévoient des dynamiques distinctes. Finalement, aux deux dernières sections, on propose deux solutions au problème de fermeture de moments qui incluent les moments d'ordre 2. Ces deux modèles de dynamiques neuronales sont originaux, et peuvent être vus comme des généralisations du modèle de Wilson–Cowan où des corrélations entre les états de différents neurones sont permises.

On verra que le premier système qu'on propose pour résoudre le problème de fermeture de moments, qui est donné aux équations (4.54), ne parvient pas à fournir une meilleure approximation macroscopique du processus stochastique sous-jacent que le modèle de champ moyen. En effet, ce modèle prévoit des solutions incohérentes avec le modèle microscopique, comme celle présentée à la figure 4.10b, ainsi que des solutions qui concordent moins

avec les trajectoires du processus stochastique que celles du modèle de champ moyen, comme l'exemple illustré aux figures 4.12 et 4.13. Cependant, le modèle présenté à la section suivante, qui est donné aux équations (4.181), mène à des résultats plus encourageants. En effet, on parvient à voir que dans certains cas, comme celui présenté à la figure 4.18, ses prédictions concordent mieux avec le modèle microscopique que celles du modèle de champ moyen.

§ 4.1. Le modèle de Wilson–Cowan

Dans cette section, on présente un bref historique des travaux qui ont mené au modèle de Wilson–Cowan, puis on donne une dérivation du modèle basée sur la publication originale de Wilson et Cowan [101]. On discute ensuite quelques généralisations de ce modèle, ainsi que les liens qu'il a avec d'autres modèles.

§ 4.1.1. Bref historique

L'étude des réseaux de neurones remonte aux travaux [72] de McCulloch et Pitts publiés en 1943 [17]. Se basant sur l'idée qu'un neurone émet un potentiel d'action s'il reçoit une excitation supérieure à un certain seuil, ces deux auteurs modélisent l'activité neuronale par un réseau de machines binaires reliées par des connexions excitatrices et inhibitrices, qui s'activent lorsqu'elles reçoivent une excitation suffisante. Ils montrent alors que n'importe quelle opération logique peut être réalisée par un tel réseau, ce qui motive l'idée d'utiliser ce genre de structure pour modéliser un cerveau.

D'après les travaux de McCulloch et Pitts, la fonction d'un réseau de neurones peut ainsi être comprise à partir de la structure des connexions entre les neurones, à la manière de portes logiques [17]. Ils cherchent alors à déterminer quelles connexions mènent à quelle opération logique. Cependant, quelques années plus tard, Shimbel et Rapoport argumentent dans [92] que cette méthode ne peut pas correspondre à la façon dont sont créés les réseaux de neurones biologiquement, puisqu'il est peu vraisemblable que des gènes commandent les détails des connexions entre les neurones. Ils soutiennent toutefois que les gènes pourraient plutôt commander les propriétés statistiques des neurones, et proposent alors d'étudier l'évolution et le fonctionnement d'un réseau de neurones d'un point de vue statistique.

Dans les décennies suivantes, plusieurs auteurs ont travaillé à décrire les réseaux de neurones selon l'approche suggérée par Shimbel et Rapoport [17], dont Beurle [3] et Griffith [39, 40]. Ces efforts ont culminé en 1972 avec les travaux de Wilson et Cowan. Ceux-ci publient dans [101] un modèle de champ moyen qui décrit, à partir d'un système dynamique, l'activité d'un réseau de neurones biologiques divisé en une population excitatrice et une population inhibitrice.

§ 4.1.2. Dérivation du modèle

Le modèle de Wilson–Cowan vise à décrire la dynamique d'un réseau de neurones biologiques d'un point de vue macroscopique en le divisant en deux populations, l'une excitatrice, et l'autre inhibitrice. Essentiellement, le modèle propose un système dynamique qui décrit comment ces deux populations s'échangent de l'information. On en présente ici une courte dérivation basée sur la présentation originale de Wilson et Cowan dans [101]. Dans ce contexte, on considère un grand réseau de neurones biologiques, où chaque neurone peut s'activer lorsqu'il reçoit une excitation suffisante provenant d'autres neurones. Un neurone est dit actif lorsqu'il émet un potentiel d'action, c'est-à-dire qu'il envoie un signal électrique à d'autres neurones. Après l'émission d'un potentiel d'action, que l'on considère instantanée, on suppose qu'un neurone passe dans un état latent, appelé *réfractaire*, dans lequel il n'est plus sensible aux signaux que lui envoient d'autres neurones. Après un certain temps passé dans cet état, il redevient sensible.

On note $E(t)$ et $I(t)$ les proportions de neurones qui émettent un potentiel d'action par unité de temps à l'instant t dans les populations excitatrice et inhibitrice. On suppose que les neurones excitateurs et inhibiteurs mettent des intervalles de temps τ_E et τ_I à émettre un potentiel d'action après avoir reçu une excitation suffisante. Afin d'étudier comment évoluent E et I , on cherche d'abord des expressions pour $E(t + \tau_E)$ et $I(t + \tau_I)$. Comme les démarches sont très similaires dans les deux cas, on ne détaille que le cas des neurones excitateurs.

On suppose que la proportion de neurones excitateurs qui émettent un potentiel d'action en $t + \tau_E$ correspond à la proportion de neurones excitateurs qui sont sensibles et qui reçoivent une excitation supérieure à leur seuil à l'instant t . On utilise cette idée pour obtenir une expression de $E(t + \tau_E)$.

D'abord, en supposant que la période réfractaire dure r_E unités de temps, la proportion de neurones excitateurs qui sont réfractaires à l'instant t est donnée par $\int_{t-r_E}^t E(s) ds$, de sorte que la proportion de neurones qui sont sensibles en t soit donnée par $1 - \int_{t-r_E}^t E(s) ds$.

Ensuite, on définit une fonction de réponse $\mathcal{S}_E : \mathbb{R} \rightarrow [0, \infty)$ afin de représenter la proportion de neurones excitateurs qui reçoivent au moins l'excitation de seuil par unité de temps. On choisit \mathcal{S}_E telle que si chaque neurone de la population excitatrice reçoit une activité constante x et que tous ses neurones sont sensibles et prêts à émettre des potentiels d'action, la proportion de ses neurones qui reçoivent au moins l'excitation de seuil par unité de temps soit donnée par $\mathcal{S}_E(x)$. La forme de \mathcal{S}_E peut être dérivée de différentes façons. Par exemple, si les seuils d'activation des neurones de la population suivent une loi de probabilité D , alors \mathcal{S}_E pourrait être donnée par

$$\mathcal{S}_E(x) = \int_{-\infty}^x D(\theta) d\theta. \quad (4.1)$$

Si la loi des seuils est unimodale, on s'attend donc à ce que \mathcal{S}_E soit de forme sigmoïdale.

Cette fonction \mathcal{S}_E donnera ainsi la proportion de neurones excitateurs qui reçoivent au moins l'excitation de seuil par unité de temps si l'on obtient une expression pour l'activité moyenne reçue par un neurone. On suppose alors que l'effet d'une excitation, chez un neurone, décroît selon une certaine fonction du temps α . L'activité moyenne reçue par un neurone excitateur au temps t est alors donnée par

$$\int_{-\infty}^t \alpha(t-s) (n_{EE}E(s) - n_{EI}I(s) + Q_E(s)) ds, \quad (4.2)$$

où $n_{JK} > 0$ représente le nombre moyen de synapses de neurones de la population K vers des neurones de la population J , et où $Q_E(t)$ est une activité entrante à la population excitatrice, mais externe au réseau.

D'après les arguments précédents, on obtient ainsi que

$$E(t + \tau_E) = \left(1 - \int_{t-r_E}^t E(s) ds\right) \mathcal{S}_E \left(\int_{-\infty}^t \alpha(t-s) (n_{EE}E(s) - n_{EI}I(s) + Q_E(s)) ds \right). \quad (4.3)$$

On peut ensuite obtenir une équation différentielle qui décrit l'évolution moyenne de E à partir de (4.3). Pour ce faire, on introduit la moyenne temporelle $\bar{E}(t) := \frac{1}{r_E} \int_{t-r_E}^t E(s) ds$. En supposant alors que $\alpha(t)$ est près de 1 pour $t \in [0, r_E]$ et décroît rapidement vers 0 lorsque $t > r_E$, on approxime $\int_{-\infty}^t \alpha(t-s) E(s) ds \approx k_E \bar{E}(t)$ pour un certain $k_E > 0$. Puisque ces moyennages temporels ont pour effet de négliger les petites fluctuations, il est cohérent de remplacer $E(t + \tau_E)$ par une approximation de Taylor de premier ordre, comme $E(t + \tau_E) \approx \bar{E}(t) + \tau_E \dot{\bar{E}}(t)$. Avec ces approximations, (4.3) devient

$$\tau_E \dot{\bar{E}}(t) = -\bar{E}(t) + (1 - r_E \bar{E}(t)) \mathcal{S}_E (k_E n_{EE} \bar{E}(t) - k_E n_{EI} \bar{I}(t) + Q_E(t)), \quad (4.4)$$

où Q_E est redéfinie par le moyennage temporel.

Afin d'alléger la notation, on note ensuite $w_{EE} := k_E n_{EE}$ et $w_{EI} := -k_E n_{EI}$. En répétant pour la population inhibitrice le travail fait pour la population excitatrice, on obtient le système

$$\tau_E \dot{\bar{E}} = -\bar{E} + (1 - r_E \bar{E}) \mathcal{S}_E (w_{EE} \bar{E} + w_{EI} \bar{I} + Q_E), \quad (4.5a)$$

$$\tau_I \dot{\bar{I}} = -\bar{I} + (1 - r_I \bar{I}) \mathcal{S}_I (w_{IE} \bar{E} + w_{II} \bar{I} + Q_I). \quad (4.5b)$$

Ces équations sont celles proposées à l'origine par Wilson et Cowan.

Malgré leur simplicité, les équations de Wilson–Cowan prévoient une dynamique assez riche. En effet, par une analyse des isoclines de pente nulle, Wilson et Cowan montrent que le système (4.5) possède entre un et cinq points fixes et qu'il peut y exister un cycle limite, selon les valeurs des paramètres. Des analyses de bifurcations ont aussi été présentées dans [29, 30, 31].

Le modèle de Wilson–Cowan a eu énormément de succès [10, 17, 21, 24], principalement parce qu'il permet de modéliser différents phénomènes neuroaux qui étaient déjà observés à l'époque, notamment la formation d'oscillations cérébrales. Il est aujourd'hui un modèle classique en neurosciences, présenté dans des manuels de référence comme [32, 49].

§ 4.1.3. Généralisations

Étant un modèle de champ moyen, le modèle de Wilson–Cowan néglige toute corrélation entre les activations de différents neurones. Or, de telles corrélations peuvent être importantes, particulièrement pour modéliser des réseaux de neurones biologiques réels [8], qui ont une taille finie. Depuis la publication du modèle, beaucoup d'efforts ont été mis en recherche pour tenter de le généraliser afin de considérer de telles corrélations. Par exemple, pour simuler du bruit, certains auteurs comme [33, 38] étendent les équations différentielles sous la forme d'équations de Langevin. Une autre approche, développée particulièrement par Buice, Cowan et Chow dans [12, 13, 14], est d'utiliser des intégrales fonctionnelles. Celles-ci permettent ensuite d'obtenir, dans certains cas, des équations différentielles pour la variance de l'activité d'un réseau [14, 19].

§ 4.1.4. Liens avec d'autres modèles

Le modèle de Wilson–Cowan se trouve dans la famille des modèles à taux de décharge¹, une famille importante de modèles en neurosciences [30, 47, 97]. Ces modèles sont également semblables au modèle proposé par Hopfield au début des années 1980 [48], qui est particulièrement populaire auprès des physiciens, ainsi qu'à celui que Grossberg a proposé au début des années 1970 [42].

Les modèles à taux de décharge sont également très importants en informatique théorique, puisque le théorème d'approximation de Funahashi et Nakamura [35] montre qu'un réseau de neurones régi par une telle dynamique peut approximer la trajectoire d'un système dynamique sur un intervalle de temps fini avec une précision arbitrairement bonne.

§ 4.2. Modélisation de l'activité d'un réseau de neurones biologiques par une dynamique MTC

L'une des limites importantes du modèle de Wilson–Cowan est qu'il ne suppose pas de corrélations entre les activations de différents neurones du réseau. Or, à partir des arguments originaux de Wilson et Cowan présentés à la section 4.1.2, il n'est pas clair de voir comment étendre leur modèle de façon à permettre de telles corrélations. Il est donc justifié de vouloir reformuler leur modèle afin de voir clairement comment en introduire.

Plusieurs auteurs ont déjà travaillé à décrire des dynamiques neuronales à partir de processus stochastiques. Parmi les premiers exemples, on trouve les travaux de Little publiés en 1974 dans [68]. Plus récemment, plusieurs auteurs ont travaillé à reformuler le modèle de Wilson–Cowan comme un processus stochastique. Dans le contexte d'un réseau de neu-

1. Traduction libre de « *firing rate models* », l'expression généralement utilisée dans la littérature anglophone.

rones biologiques divisé en n populations, de nombreux auteurs, dont [2, 9, 12, 14, 21, 28], ont proposé des équations de Kolmogorov afin de décrire la probabilité que n_j neurones soient actifs dans la j -ième population de sorte que l'approximation de champ moyen de ces équations corresponde au modèle de Wilson–Cowan. Plusieurs auteurs, dont [9, 12, 14], proposent alors d'utiliser ces équations de Kolmogorov afin d'étudier l'évolution des seconds moments à partir d'intégrales fonctionnelles. Ces modèles considèrent seulement deux états possibles aux neurones, qui décrivent simplement si un neurone est en train ou non d'émettre un signal. Or, biologiquement, un neurone peut aussi se trouver dans un état réfractaire.

On propose ici un nouveau modèle pour représenter une dynamique neuronale par un processus de Markov qui permet trois états aux neurones. Pour ce faire, on propose de reformuler le modèle de Wilson–Cowan à partir d'une dynamique MTC dont l'approximation de champ moyen permet de retrouver les équations (4.5). Cette méthode a l'avantage de fournir une façon claire d'introduire des corrélations entre les états de différents neurones. En effet, le problème de fermeture de moments surviendra de la même façon que dans le cas du modèle SIR abordé au chapitre 3. On verra alors aux sections suivantes différentes fermetures de moments qui permettent d'inclure ou non les covariances.

On note que des chaînes de Markov ont déjà été proposées pour modéliser les états de neurones en permettant trois états. Par exemple, Cowan a proposé un tel modèle dans [20], mais sans analyse approfondie du cas à trois états. De plus, le modèle présenté par Zare-pour *et al.* dans [104] est semblable à celui qu'on propose ici, mais les taux qui caractérisent les transitions des neurones sont différents.

§ 4.2.1. Modélisation du réseau de neurones biologiques

On considère un réseau de N neurones dont on veut représenter l'activité par une dynamique MTC définie sur un graphe aléatoire. On prend alors une matrice de poids W de dimension N et G le graphe aléatoire induit par cette matrice de poids. On interprète les nœuds de G comme les neurones, et on interprète une composante W_{jk} de la matrice de poids comme l'activité que le neurone k envoie au neurone j lorsqu'il s'active. Cette valeur est donc positive si k est excitateur, et négative si k est inhibiteur.

Afin de représenter les états des neurones, on introduit un espace d'états ternaires (E, \mathcal{E}) pour G . On interprète alors :

- l'état 0 comme celui d'un neurone sensible, prêt à réagir à une excitation,
- l'état 1 comme celui d'un neurone actif, en train d'émettre un potentiel d'action,
- l'état i comme celui d'un neurone réfractaire.

On veut ensuite définir les taux de transition a_j , b_j et c_j pour chaque neurone j du graphe de façon à reproduire la dynamique de seuil qui régit l'activation des neurones. Pour ce

faire, on introduit d'abord sur le graphe quatre paramètres d'ordre 1, notés α, β, γ et θ . On interprète les composantes de θ comme les seuils d'activation des neurones du réseau, et on interprète les composantes de α, β et γ comme les taux de transition entre les trois états pour un neurone qui recevrait toujours l'excitation de seuil, selon le diagramme donné à la figure 4.1. On supposera toujours que α, β et γ sont strictement positifs p.s.

Il est simple d'expliciter un espace des paramètres sur lequel le graphe et ses paramètres peuvent être définis. En effet, soient $\mu_W: \mathcal{B}(\mathbb{R}^{N \times N}) \rightarrow [0, 1]$ la loi de probabilité voulue pour la matrice de poids et $\mu_\alpha, \mu_\beta, \mu_\gamma, \mu_\theta: \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ les lois des paramètres d'ordre 1. Si l'on prend alors $\Theta := \mathbb{R}^{N^6}$, $\mathcal{G} := \mathcal{B}(\mathbb{R}^{N^6})$ et $\mu := \mu_\alpha \otimes \mu_\beta \otimes \mu_\gamma \otimes \mu_\theta \otimes \mu_W$, l'espace de probabilité $(\Theta, \mathcal{G}, \mu)$ forme l'espace des paramètres du graphe. Chacun des paramètres peut alors être vu comme une projection sur l'une des cinq composantes si l'on voit $\mathbb{R}^{N^6} = (\mathbb{R}^N)^4 \times \mathbb{R}^{N \times N}$. On choisit les lois de probabilité des paramètres de sorte qu'ils soient tous intégrables.

Il est alors possible de définir les taux de transition des neurones. On veut qu'un neurone j s'active au taux α_j dès qu'il reçoit une excitation supérieure à son seuil θ_j , et qu'il ne puisse pas s'activer sinon. On définit alors

$$a_j(\vartheta, x) := \alpha_j \mathbb{1}_{T_j}(\vartheta) \quad \text{où} \quad T_j := \left\{ \vartheta' \in \Theta : \sum_{k=1}^N W_{jk}(\vartheta') \operatorname{Re} x_k + Q_j > \theta_j(\vartheta') \right\} \in \mathcal{G}, \quad (4.6)$$

où $Q_j \in \mathbb{R}$ représente une activité entrante au neurone j externe au réseau, et où l'on sait que $T_j \in \mathcal{G}$ puisque les paramètres sont des projections, qui sont \mathcal{G} -mesurables. Les autres taux de transition sont plus simples à définir, puisqu'ils ne dépendent pas de l'état du graphe. Ainsi, on prend simplement

$$b_j(\vartheta, x) := \beta_j(\vartheta) \quad \text{et} \quad c_j(\vartheta, x) := \gamma_j(\vartheta). \quad (4.7)$$

Comme les paramètres sont tous \mathcal{G} -mesurables et intégrables, on voit de (4.6) et de (4.7) que pour chaque $x \in E$ les projections $a_j(\cdot, x), b_j(\cdot, x)$ et $c_j(\cdot, x)$ sont toutes \mathcal{G} -mesurables et intégrables. Ainsi, par le théorème 2.3, si P est une mesure de probabilité sur (E, \mathcal{E}) , pour tout $\vartheta \in \Theta$ il existe une mesure de probabilité \mathbb{P}^ϑ sur $(\Omega, \mathcal{F}) := (E, \mathcal{E})^{[0, \infty)}$ telle que les conditions *i*) à *iii*) du théorème sont respectées, et telle que le processus de projection de coordonnées $\{X_t\}_{t \geq 0}$ respecte les conditions (2.7) données au corollaire 1. De plus, par le corollaire 2, l'espace $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F})$ peut être muni d'une mesure de probabilité \mathbb{P} telle que pour toute variable aléatoire Z sur $\Theta \times \Omega$,

$$\mathbb{E}[Z] = \int_{\Theta} \int_{\Omega} Z(\vartheta, \omega) d\mathbb{P}^\vartheta(\omega) d\mu(\vartheta), \quad (4.8)$$

où \mathbb{E} est l'espérance par rapport à \mathbb{P} . Le processus $\{X_t\}_{t \geq 0}$ correspond alors à une dynamique MTC qui représente l'état du réseau de neurones selon le temps.

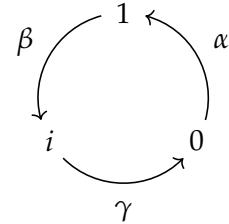


FIGURE 4.1 – Transitions entre les trois états d'un neurone.

§ 4.2.2. Passage à la description macroscopique de la dynamique

On veut maintenant décrire la dynamique sur le graphe d'un point de vue macroscopique. Pour ce faire, on divise d'abord le graphe en une collection \mathcal{P} de n populations, et on cherche à appliquer les résultats de la section 2.4. Comme on a vu que les projections des taux de transition sur l'espace des paramètres sont intégrables, on cherche à trouver des taux de transition moyens de la forme (2.59) de sorte que la démarche présentée à la section 2.4 donne une description macroscopique moyenne de la dynamique.

En reprenant la notation de la section 2.4, on cherche pour chaque population $J \in \mathcal{P}$ des fonctions $a_J, b_J, c_J : [0, 1]^{2n} \rightarrow [0, \infty)$ telles que pour tout $t \geq 0$,

$$a_J(Y_t) = \frac{1}{|J_t^0|} \sum_{j \in J_t^0} a_j(\vartheta, X_t) = \frac{1}{|J_t^0|} \sum_{j \in J_t^0} \alpha_j \mathbb{1}_{\{\sum_{k=1}^N W_{jk} \operatorname{Re} X_t^k + Q_j > \theta_j\}}(\vartheta), \quad (4.9a)$$

$$b_J(Y_t) = \frac{1}{|J_t^1|} \sum_{j \in J_t^1} b_j(\vartheta, X_t) = \frac{1}{|J_t^1|} \sum_{j \in J_t^1} \beta_j(\vartheta) \quad (4.9b)$$

et

$$c_J(Y_t) = \frac{1}{|J_t^i|} \sum_{j \in J_t^i} c_j(\vartheta, X_t) = \frac{1}{|J_t^i|} \sum_{j \in J_t^i} \gamma_j(\vartheta), \quad (4.9c)$$

où les inégalités tiennent \mathbb{P} -p.s. et où Y_t note l'état macroscopique donné par X_t selon (2.49).

Comme dans le cas du modèle SIR présenté au chapitre 3, il est difficile de trouver de telles fonctions dans le cas général. On trouvera donc plutôt des fonctions qui devraient satisfaire ces contraintes approximativement lorsque les populations du graphe sont très grandes. Pour ce faire, on suppose que les paramètres sont homogènes sur \mathcal{P} au sens de la définition 2.7, afin de reprendre les idées discutées à la section 3.2.3. On suppose également que les activités entrantes externes Q_j sont constantes à travers chaque population. On notera donc ces activités entrantes externes avec des indices qui renvoient aux populations plutôt qu'aux neurones.

En se basant sur la loi des grands nombres et en supposant que les ensembles J_t^1 et J_t^i sont très grands avec une probabilité élevée, on approxime que

$$\frac{1}{|J_t^1|} \sum_{j \in J_t^1} \beta_j \approx \mathbb{E}[\beta_j] =: \beta_J \quad \text{et que} \quad \frac{1}{|J_t^i|} \sum_{j \in J_t^i} \gamma_j \approx \mathbb{E}[\gamma_j] =: \gamma_J. \quad (4.10)$$

De la même façon, on approxime que

$$\sum_{k=1}^N W_{jk} \operatorname{Re} X_t^k = \sum_{K \in \mathcal{P}} \sum_{k \in K^1} W_{jk} \approx \sum_{K \in \mathcal{P}} |K_t^1| W_{JK} = \sum_{K \in \mathcal{P}} |K| W_{JK} A_t^K, \quad (4.11)$$

où $W_{JK} := \mathbb{E}[W_{jk}]$ pour $j \in J$ et $k \in K$. Afin d'alléger la notation pour la suite, on définit $c_{JK} := |K| W_{JK}$, un poids de connexion qui représente l'activité que la population K envoie

à la population J si tous ses neurones sont actifs. On définit également

$$B_t^J := \sum_{K \in \mathcal{P}} c_{JK} A_t^K + Q_J, \quad (4.12)$$

où Q_J est l'activité entrante externe à la population J . Ainsi, B_t^J représente l'activité totale entrante dans la population J au temps t .

En utilisant (4.11) et en suivant le même principe que dans les derniers cas,

$$\frac{1}{|J_t^0|} \sum_{j \in J_t^0} \alpha_j \mathbb{1}_{\left\{ \sum_{k=1}^N W_{jk} \operatorname{Re} X_t^k + Q_J > \theta_j \right\}} \approx \frac{1}{|J_t^0|} \sum_{j \in J_t^0} \alpha_j \mathbb{1}_{\{B_t^J > \theta_j\}} \approx \mathbb{E}[\alpha_j \mathbb{1}_{\{B_t^J > \theta_j\}}]. \quad (4.13)$$

Il est important de rappeler ici que l'intuition sur la seconde approximation est basée sur la loi des grands nombres. Elle repose sur l'idée que l'ensemble sur lequel les variables aléatoires sont sommées est très grand avec une probabilité très élevée, et que la composition de cet ensemble est indépendante des variables aléatoires sommées. Or, l'intuition suggère qu'il pourrait y avoir une corrélation entre l'activation d'un neurone et son seuil : si un neurone a un seuil très élevé, il risque moins de recevoir suffisamment d'excitation pour s'activer. Ainsi, on peut s'attendre à ce que la seconde approximation en (4.13) soit de moins en moins valide lorsque la variance des seuils est de plus en plus élevée.

L'approximation (4.13) peut être réécrite en utilisant le fait que les paramètres α_j et θ_j sont indépendants et en négligeant une possible corrélation entre α_j et B_t^J . En effet, ceci donne que

$$\mathbb{E}[\alpha_j \mathbb{1}_{\{B_t^J > \theta_j\}}] \approx \mathbb{E}[\alpha_j] \mathbb{E}[\mathbb{1}_{\{B_t^J > \theta_j\}}] = \mathbb{E}[\alpha_j] \mathbb{P}[B_t^J > \theta_j] = \alpha_J F_{\theta_J}(B_t^J), \quad (4.14)$$

où $\alpha_J := \mathbb{E}[\alpha_j]$ et F_{θ_J} note la fonction de répartition des seuils θ_j pour $j \in J$.

En se basant sur (4.10), ainsi que sur (4.13) et (4.14), on définit donc les taux de transition moyens

$$a_J(Y_t) := \alpha_J F_{\theta_J}(B_t^J), \quad b_J(Y_t) := \beta_J \quad \text{et} \quad c_J(Y_t) := \gamma_J, \quad (4.15)$$

où B_t^J est donné en fonction des composantes de l'état macroscopique Y_t selon (4.12). On remarque que, puisque les paramètres α , β et γ sont strictement positifs p.s., alors $\alpha_J, \beta_J, \gamma_J > 0$. Avec ces taux de transition, le système (2.60) donné à la section (2.4) prend la forme

$$\dot{\mathcal{A}}_J(t) = -\beta_J \mathcal{A}_J(t) + \alpha_J \mathbb{E}[S_t^J F_{\theta_J}(B_t^J)], \quad (4.16a)$$

$$\dot{\mathcal{R}}_J(t) = -\gamma_J \mathcal{R}_J(t) + \beta_J \mathcal{A}_J(t), \quad (4.16b)$$

$$\dot{\mathcal{S}}_J(t) = -\alpha_J \mathbb{E}[S_t^J F_{\theta_J}(B_t^J)] + \gamma_J \mathcal{R}_J(t). \quad (4.16c)$$

Il reste ensuite à rendre ce système autonome en trouvant une solution au problème de fermeture de moments.

Naturellement, la première façon d'obtenir un système autonome à partir de (4.16) est de négliger les variances des fractions de populations afin d'en obtenir l'approximation de

champ moyen. On approxime alors que $\mathbb{E}[S_t^J F_{\theta_J}(B_t^J)] = S_t^J(t) \mathbb{E}[F_{\theta_J}(B_t^J)] = S_t^J(t) F_{\theta_J}(\mathcal{B}_J(t))$, où $\mathcal{B}_J(t) := \mathbb{E}[B_t^J]$. Ainsi, l'approximation de champ moyen de (4.16) est le système

$$\dot{\mathcal{A}}_J = -\beta_J \mathcal{A}_J + \alpha_J S_J F_{\theta_J}(\mathcal{B}_J), \quad (4.17a)$$

$$\dot{\mathcal{R}}_J = -\gamma_J \mathcal{R}_J + \beta_J \mathcal{A}_J, \quad (4.17b)$$

$$\dot{\mathcal{S}}_J = -\alpha_J S_J F_{\theta_J}(\mathcal{B}_J) + \gamma_J \mathcal{R}_J. \quad (4.17c)$$

§ 4.2.3. Réduction au modèle de Wilson–Cowan

La correspondance entre l'approximation de champ moyen (4.17) du système (4.16) et les équations (4.5) de Wilson–Cowan n'est pas complètement directe. En effet, le modèle de Wilson–Cowan ne considère pas explicitement que les fractions réfractaires des populations excitatrice et inhibitrice varient dans le temps. En reprenant la notation de la section 4.1.2, les proportions de neurones excitateurs et inhibiteurs qui sont réfractaires en un temps t sont données respectivement par $\int_{t-r_E}^t E(s) ds = r_E \bar{E}(t)$ et par $\int_{t-r_I}^t I(s) ds = r_I \bar{I}(t)$. Ainsi, dans le contexte du modèle de Wilson–Cowan, la fraction réfractaire d'une population de neurones est toujours vue comme proportionnelle à sa fraction active. Or, cette supposition n'est pas faite dans le contexte du système (4.17).

Pour retrouver le modèle de Wilson–Cowan à partir de (4.17), il est donc logique de fixer les fractions réfractaires des populations à leurs solutions d'équilibre. En imposant $\dot{\mathcal{R}}_J = 0$ dans (4.17b), on obtient que

$$0 = -\gamma_J \mathcal{R}_J + \beta_J \mathcal{A}_J \implies \mathcal{R}_J = \frac{\beta_J}{\gamma_J} \mathcal{A}_J. \quad (4.18)$$

En utilisant l'équivalence $S_J = 1 - \mathcal{A}_J - \mathcal{R}_J$ et en utilisant la solution d'équilibre (4.18) dans l'expression (4.17a) de $\dot{\mathcal{A}}_J$, on trouve que

$$\dot{\mathcal{A}}_J = -\beta_J \mathcal{A}_J + \alpha_J \left(1 - \left(1 + \frac{\beta_J}{\gamma_J}\right) \mathcal{A}_J\right) F_{\theta_J}(\mathcal{B}_J). \quad (4.19)$$

La forme de cette équation est presque la même que celle de (4.5). Cependant, les paramètres ne sont pas directement équivalents. De plus, les variables dynamiques \bar{E} et \bar{I} des équations de Wilson–Cowan ne sont pas des fractions actives de populations, mais plutôt des taux d'activation. Ainsi, pour comprendre l'équivalence entre les deux modèles, il convient de définir des taux à partir des activités \mathcal{A}_J . On définit donc $\mathcal{T}_J := \beta_J \mathcal{A}_J$ pour chaque $J \in \mathcal{P}$. En substituant dans (4.19), on trouve alors que

$$\frac{1}{\beta_J} \dot{\mathcal{T}}_J = -\mathcal{T}_J + \alpha_J \left(1 - \left(\frac{1}{\beta_J} + \frac{1}{\gamma_J}\right) \mathcal{T}_J\right) F_{\theta_J}(\mathcal{B}_J). \quad (4.20)$$

Cette équation correspond alors exactement à ce qu'ont proposé Wilson et Cowan, mais pour un nombre quelconque de populations. En effet, dans le cas où le graphe est divisé en

deux populations notées E et I , on obtient le système

$$\frac{1}{\beta_E} \dot{\mathcal{T}}_E = -\mathcal{T}_E + \left(1 - \left(\frac{1}{\beta_E} + \frac{1}{\gamma_E}\right) \mathcal{T}_E\right) \alpha_E F_{\theta_E} \left(\frac{c_{EE}}{\beta_E} \mathcal{T}_E + \frac{c_{EI}}{\beta_I} \mathcal{T}_I + Q_E\right) \quad (4.21a)$$

$$\frac{1}{\beta_I} \dot{\mathcal{T}}_I = -\mathcal{T}_I + \left(1 - \left(\frac{1}{\beta_I} + \frac{1}{\gamma_I}\right) \mathcal{T}_I\right) \alpha_I F_{\theta_I} \left(\frac{c_{IE}}{\beta_E} \mathcal{T}_E + \frac{c_{II}}{\beta_I} \mathcal{T}_I + Q_I\right). \quad (4.21b)$$

On peut alors voir les équivalences suivantes entre les quantités utilisées par Wilson et Cowan et celles qu'on propose ici :

$$\tau_J = \frac{1}{\beta_J}, \quad r_J = \frac{1}{\beta_J} + \frac{1}{\gamma_J}, \quad w_{JK} = \frac{c_{JK}}{\beta_K}, \quad \text{et} \quad \delta_J = \alpha_J F_{\theta_J}. \quad (4.22)$$

Ces équivalences font bien en sorte que les unités de ces quantités sont cohérentes entre elles, et en les substituant dans (4.21), on retrouve bien les équations classiques (4.5).

§ 4.3. Effet de l'état réfractaire sur le modèle de Wilson–Cowan

On a vu à la section précédente une façon de modéliser la dynamique d'un réseau de neurones biologiques par une dynamique MTC sur un graphe aléatoire. On a aussi vu que le modèle de Wilson–Cowan pouvait être obtenu à partir de l'approximation de champ moyen de ce modèle, en fixant les fractions réfractaires des populations à leurs solutions d'équilibre. Le modèle de Wilson–Cowan néglige donc une partie de la dynamique considérée par l'approximation de champ moyen du modèle construit à la section précédente. Il est alors légitime de se demander si des aspects importants de la dynamique sont perdus par cette simplification.

Bien qu'une période réfractaire soit considérée dans la version originale du modèle de Wilson–Cowan, plusieurs auteurs ont choisi de la négliger et la fixent à zéro [22, 32]. Or, plus récemment, d'autres auteurs dont [87, 99] ont argumenté que la période réfractaire avait un réel effet sur la dynamique d'un réseau de neurones biologiques. Weistuch *et al.* ont d'ailleurs présenté dans [99] un modèle similaire à celui qu'on présente ici, mais en temps discret et obtenu différemment.

Dans cette section, on s'intéresse d'abord à savoir si le modèle de Wilson–Cowan permet de reproduire le comportement des solutions à long terme de l'approximation de champ moyen du modèle construit à la section précédente. On veut ainsi vérifier s'il est valide de fixer d'emblée les fractions réfractaires des populations à leurs solutions d'équilibre. Après avoir donné quelques résultats généraux, on étudiera deux exemples où les deux modèles prévoient des solutions d'équilibre très différentes.

§ 4.3.1. Quelques résultats généraux

On veut ici comparer le système (4.17), que l'on peut réécrire comme

$$\dot{\mathcal{A}}_J = -\beta_J \mathcal{A}_J + \alpha_J (1 - \mathcal{A}_J - \mathcal{R}_J) F_{\theta_J}(\mathcal{B}_J), \quad (4.17a)$$

$$\dot{\mathcal{R}}_J = -\gamma_J \mathcal{R}_J + \beta_J \mathcal{A}_J, \quad (4.17b)$$

au système

$$\dot{\mathcal{A}}_J = -\beta_J \mathcal{A}_J + \alpha_J \left(1 - \left(1 + \frac{\beta_J}{\gamma_J} \right) \mathcal{A}_J \right) F_{\theta_J}(\mathcal{B}_J). \quad (4.19)$$

Le système (4.17) est défini dans \mathbb{R}^{2n} , puisqu'il y a deux axes pour chacune des n populations, alors que le système (4.19) est défini dans \mathbb{R}^n , puisqu'il n'y a qu'un axe par population. Cependant, ce ne sont pas tous les points de \mathbb{R}^{2n} qui correspondent à des états biologiques du réseau de neurones. En effet, pour chaque population J , les quantités \mathcal{A}_J , \mathcal{R}_J et $\mathcal{S}_J = 1 - \mathcal{A}_J - \mathcal{R}_J$ représentent les espérances de fractions des neurones qui composent J . Elles doivent donc toutes être entre 0 et 1. Ainsi, seuls les points de l'ensemble

$$\mathcal{D}_n := \{ (\mathcal{A}_J, \mathcal{R}_J)_{J \in \mathcal{P}} \in [0, 1]^{2n} : \forall J \in \mathcal{P}, \mathcal{A}_J + \mathcal{R}_J \leq 1 \} \quad (4.23)$$

correspondent vraiment à des états macroscopiques du réseau de neurones.

Pour l'analyse de ces systèmes, on suppose que les F_{θ_J} sont de classe \mathcal{C}^1 sur \mathbb{R} . Dans ce cas, les champs vectoriels associés aux systèmes (4.17) et (4.19) sont tous deux de classe \mathcal{C}^1 , et sont donc localement Lipschitz. Par le théorème 1.2, les deux systèmes admettent donc des solutions uniques pour toutes conditions initiales dans leurs domaines respectifs.

Comme première étape de comparaison entre les deux systèmes, on peut remarquer une correspondance entre leurs points fixes.

Proposition 4.1. *Un point $(\mathcal{A}_J^*)_{J \in \mathcal{P}} \in \mathbb{R}^n$ est un point fixe du système (4.19) si et seulement si le point $(\mathcal{A}_J^*, \frac{\beta_J}{\gamma_J} \mathcal{A}_J^*)_{J \in \mathcal{P}} \in \mathbb{R}^{2n}$ est un point fixe du système (4.17).* ◆

Démonstration. Soient $(\mathcal{A}_J^*)_{J \in \mathcal{P}}$ un point fixe du système (4.19) et $\mathcal{R}_J^* := \frac{\beta_J}{\gamma_J} \mathcal{A}_J^*$ pour $J \in \mathcal{P}$. D'après (4.17a), pour tout $J \in \mathcal{P}$,

$$\dot{\mathcal{A}}_J \Big|_{(\mathcal{A}_J, \mathcal{R}_J) = (\mathcal{A}_J^*, \mathcal{R}_J^*)} = -\beta_J \mathcal{A}_J^* + \alpha_J \left(1 - \mathcal{A}_J^* - \frac{\beta_J}{\gamma_J} \mathcal{A}_J^* \right) F_{\theta_J}(\mathcal{B}_J) = 0, \quad (4.24)$$

puisque $(\mathcal{A}_J^*)_{J \in \mathcal{P}}$ est un point fixe de (4.19). Or,

$$\dot{\mathcal{R}}_J \Big|_{(\mathcal{A}_J, \mathcal{R}_J) = (\mathcal{A}_J^*, \mathcal{R}_J^*)} = -\gamma_J \frac{\beta_J}{\gamma_J} \mathcal{A}_J^* + \beta_J \mathcal{A}_J^* = 0. \quad (4.25)$$

Ainsi, le point $(\mathcal{A}_J^*, \mathcal{R}_J^*)_{J \in \mathcal{P}}$ est un point fixe de (4.17). De la même façon, si $(\mathcal{A}_J^*, \frac{\beta_J}{\gamma_J} \mathcal{A}_J^*)_{J \in \mathcal{P}}$ est un point fixe de (4.17), alors l'évaluation de la dérivée $\dot{\mathcal{A}}_J$ donnée par (4.19) au point $\mathcal{A}_J = \mathcal{A}_J^*$ doit être nulle, puisque la substitution de $(\mathcal{A}_J, \mathcal{R}_J) = (\mathcal{A}_J^*, \frac{\beta_J}{\gamma_J} \mathcal{A}_J^*)$ dans (4.17a) doit être nulle. ■

Il n'est pas clair, toutefois, si la stabilité des points fixes des deux systèmes peut être reliée.

Pour aller plus loin, il est intéressant de noter que le système (4.19) peut être vu comme un sous-espace du système (4.17). En effet, les points où les fractions réfractaires sont fixées à leurs solutions d'équilibre forment un sous-espace euclidien de dimension n dans \mathbb{R}^{2n} , sur lequel les deux équations différentielles donnent les mêmes champs vectoriels. L'intersection de ce sous-espace euclidien avec le domaine \mathcal{D}_n est simplement donnée, pour n populations, par

$$\mathcal{D}_n^{\text{wc}} := \left\{ (\mathcal{A}_J, \mathcal{R}_J)_{J \in \mathcal{P}} \in \mathcal{D}_n : \forall J \in \mathcal{P}, \mathcal{R}_J = \frac{\beta_J}{\gamma_J} \mathcal{A}_J \right\}. \quad (4.26)$$

Il est alors intuitivement clair que le passage du système (4.17) au système (4.19) ne néglige pas d'information importante sur la dynamique du réseau de neurones si l'ensemble $\mathcal{D}_n^{\text{wc}}$ est attractif et si son bassin d'attraction inclus le domaine \mathcal{D}_n . Malheureusement, il n'est pas difficile de voir que ce n'est pas le cas.

| **Proposition 4.2.** *L'ensemble $\mathcal{D}_n^{\text{wc}}$ n'est pas invariant par le flot du système dynamique (4.17).* ◆

Démonstration. L'ensemble $\mathcal{D}_n^{\text{wc}}$ est une région d'un sous-espace euclidien de dimension n dans \mathbb{R}^{2n} . D'après le théorème 1.4, pour que cet ensemble soit invariant par le flot du système (4.17), il est donc nécessaire que le champ vectoriel donné par l'équation différentielle (4.17) soit tangent à $\mathcal{D}_n^{\text{wc}}$ en chacun de ses points. Or, il est impossible que ce soit le cas, puisque le champ vectoriel n'a pas de composante parallèle aux axes associés aux composantes \mathcal{R}_J de l'état macroscopique, mais $\mathcal{D}_n^{\text{wc}}$ n'est pas orthogonal à ces axes.

Pour le voir plus clairement, on fixe $J \in \mathcal{P}$ et $x \in \mathbb{R}$ tel que

$$0 < \frac{\alpha_J \gamma_J}{\alpha_J \beta_J + \alpha_J \gamma_J + \beta_J \gamma_J} < x < \frac{\alpha_J \gamma_J}{\alpha_J \beta_J + \alpha_J \gamma_J + 0} = \frac{\gamma_J}{\beta_J + \gamma_J} < 1. \quad (4.27)$$

Soit $Y \in \mathcal{D}_n^{\text{wc}}$ un point dont la composante selon l'axe \mathcal{A}_J vaut x , la composante selon l'axe \mathcal{R}_J vaut $\frac{\beta_J}{\gamma_J}x$, et les composantes selon les autres axes sont nulles. On peut alors remarquer que le vecteur Y^\perp avec 1 selon l'axe \mathcal{A}_J , $-\gamma_J/\beta_J$ selon l'axe \mathcal{R}_J et 0 selon les autres axes est orthogonal à $\mathcal{D}_n^{\text{wc}}$ en Y . Comme $\dot{\mathcal{R}}_J|_Y = 0$, alors le produit scalaire de Y^\perp par la valeur du champ vectoriel associé à (4.17) en Y vaut simplement

$$\dot{\mathcal{A}}_J \Big|_Y = -\beta_J x + \alpha_J \left(1 - \frac{\beta_J + \gamma_J}{\gamma_J} x\right) F_{\theta_J}(c_{JJ} x + Q_J) \quad (4.28)$$

$$\leq -\beta_J x + \alpha_J \left(1 - \frac{\beta_J + \gamma_J}{\gamma_J} x\right) \quad (4.29)$$

$$= \alpha_J - \frac{\beta_J \gamma_J + \alpha_J \beta_J + \alpha_J \gamma_J}{\gamma_J} < \alpha_J - \alpha_J = 0, \quad (4.30)$$

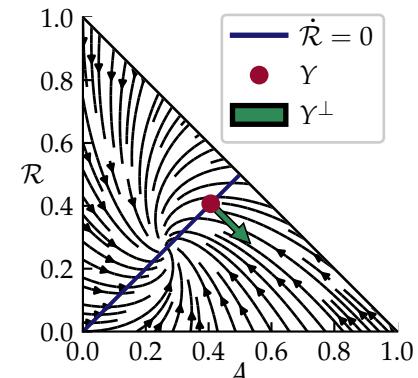


FIGURE 4.2 – Exemple de portrait de phase du système (4.17) pour une seule population. L'isocline $\dot{\mathcal{R}} = 0$ est l'ensemble $\mathcal{D}_1^{\text{wc}}$.

où l'inégalité (4.29) suit du fait que $F_{\theta_j} \geq 0$ puisqu'elle est une fonction de répartition, et où l'inégalité (4.30) suit de (4.27). Ce raisonnement est illustré à la figure 4.2 pour le cas d'une seule population. On conclut alors que le champ vectoriel associé à (4.17) n'est pas tangent à $\mathcal{D}_n^{\text{WC}}$ en au moins un de ses points. Par le théorème 1.4, cet ensemble ne peut donc pas être invariant par le flot de (4.17). ■

On sait toutefois qu'un ensemble ne peut être attractif que s'il est invariant. Le résultat précédent implique alors directement le suivant.

| **Corollaire.** *L'ensemble $\mathcal{D}_n^{\text{WC}}$ n'est pas un ensemble attractif du système dynamique (4.17).* ◆

Il n'est donc pas possible de conclure qu'en général, les solutions d'équilibre du système (4.19) soient une bonne approximation de celles du système (4.17).

§ 4.3.2. Étude d'un exemple à une population

Étant donné leur caractère non linéaire, il est difficile d'étudier les différences entre les systèmes dynamiques (4.17) et (4.19) d'un point de vue général. Cependant, on peut trouver des exemples où les deux modèles prévoient des comportements différents. On s'intéresse dans cette sous-section à un exemple particulier où les comportements à long terme des solutions prévues par (4.17) et (4.19) sont qualitativement très différentes, même en prenant dans les deux cas la même condition initiale.

On considère un cas où le graphe n'est composé que d'une population. Pour simplifier la notation, on laissera donc tomber l'indice qui réfèrerait à cette unique population. On suppose que les seuils suivent une loi logistique dans la population, de sorte que

$$F_\theta(y) = \frac{1}{1 + \exp\left(\frac{y-\theta}{s_\theta}\right)}, \quad (4.31a)$$

où s_θ est un facteur d'échelle qui correspond à l'écart type des seuils à un facteur $\pi/\sqrt{3}$ près. On fixe également les paramètres

$$\alpha = 5 \text{ ms}^{-1}, \quad \beta = 6/5 \text{ ms}^{-1}, \quad \gamma = 2/5 \text{ ms}^{-1}, \quad \theta = 2, \quad s_\theta = 2/5, \quad Q = 0, \quad c = 8. \quad (4.31b)$$

On choisit ensuite l'état initial

$$\mathcal{Y}_0 = (\mathcal{A}_0, \mathcal{R}_0) := (1/10, 3/10). \quad (4.32)$$

On remarque que

$$\mathcal{R}_0 = \frac{3}{10} = \frac{6}{5} \cdot \frac{5}{2} \cdot \frac{1}{10} = \frac{\beta}{\gamma} \mathcal{A}_0, \quad (4.33)$$

de sorte que la composante \mathcal{R} de l'état initial (4.32) se trouve déjà à sa solution d'équilibre, c'est-à-dire que $\mathcal{Y}_0 \in \mathcal{D}_1^{\text{WC}}$.

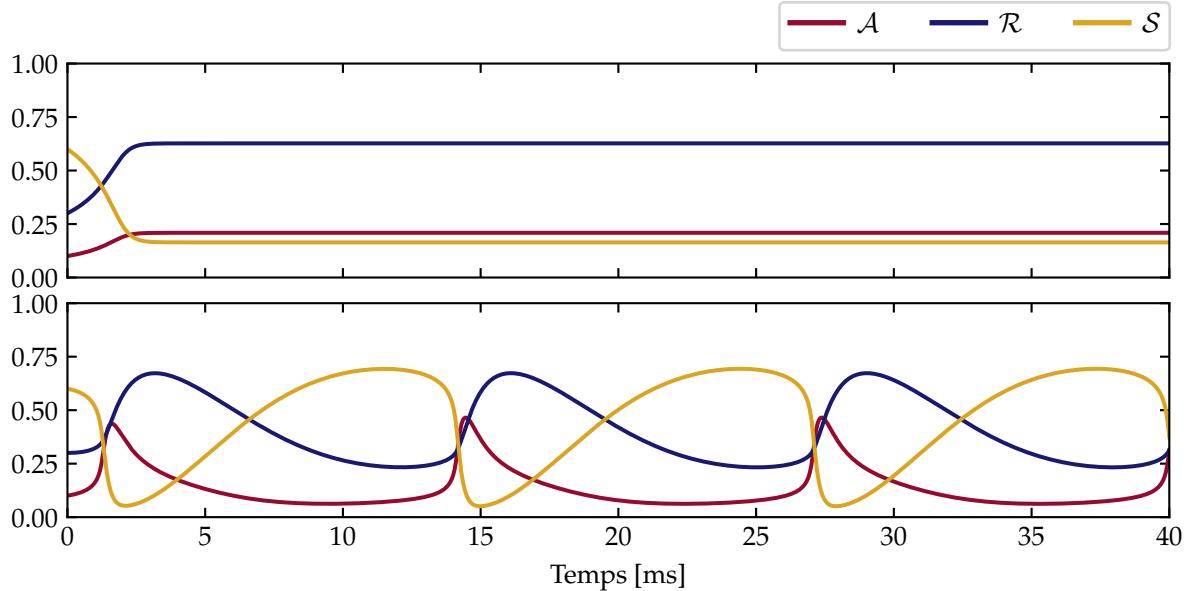


FIGURE 4.3 – En haut, solution de (4.19) pour les paramètres (4.31) et la condition initiale (4.32). En bas, solution de (4.17) pour les mêmes paramètres et la même condition initiale.

Les solutions des systèmes (4.17) et (4.19) peuvent être obtenues numériquement pour les paramètres (4.31) à partir de la condition initiale (4.32). Les résultats sont présentés à la figure 4.3. On y voit que le modèle où les fractions réfractaires sont fixées d'avance en fonction des fractions actives prévoit que la solution converge vers un point fixe stable, alors que le modèle de champ moyen prévoit que la solution converge vers un cycle limite. Celui-ci peut être visualisé plus facilement sur un portrait de phase du système de champ moyen. Un tel portrait de phase est donné à la figure 4.4.

Ce résultat est particulièrement intéressant, puisque le modèle (4.19) où les fractions réfractaires sont fixées n'a ici qu'une dimension, puisqu'il n'y a qu'une population. Ainsi, il ne pourrait en aucun cas prévoir de cycle limite. De plus, d'après la proposition 4.1, le point fixe vers lequel la solution de (4.19) converge doit aussi être un point fixe du système (4.17), mais il semble ne plus être stable.

Pour mieux comprendre ce qui se produit, on peut combiner les deux systèmes en un seul

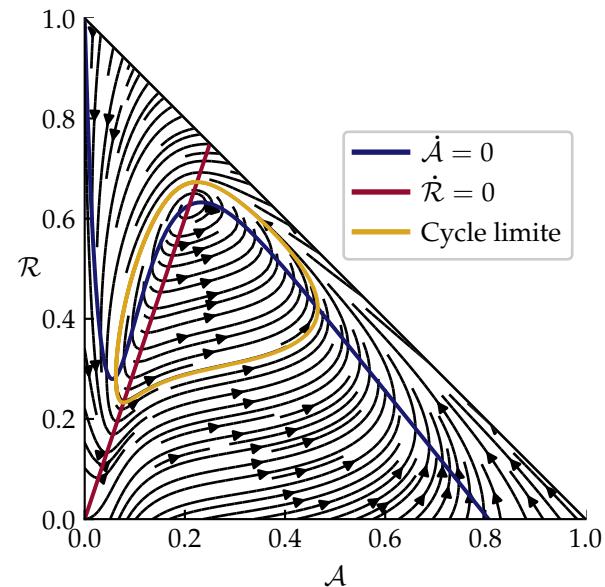


FIGURE 4.4 – Portrait de phase du système de champ moyen (4.17) pour les paramètres (4.31).

en ajoutant un paramètre supplémentaire, qu'on note ε , afin de faire une transition entre les deux systèmes. Pour ce faire, on considère le système

$$\dot{\mathcal{A}} = -\beta\mathcal{A} + \alpha(1 - \mathcal{A} - \mathcal{R})F_\theta(\mathcal{B}), \quad (4.34a)$$

$$\varepsilon\dot{\mathcal{R}} = -\gamma\mathcal{R} + \beta\mathcal{A}. \quad (4.34b)$$

Le cas $\varepsilon = 0$ correspond ainsi à fixer \mathcal{R} à sa solution d'équilibre et revient au système (4.19), alors que le cas $\varepsilon = 1$ correspond au système de champ moyen (4.17).

Il est alors possible d'étudier ce qui se produit entre les deux systèmes lorsque la valeur du paramètre ε varie entre 0 et 1. Pour ce faire, on calcule d'abord la matrice jacobienne du système (4.34). On obtient la matrice

$$\mathcal{J} = \begin{pmatrix} -\beta + \alpha F_\theta(\mathcal{B}) + \alpha \delta F'_\theta(\mathcal{B})c & -\alpha F_\theta(\mathcal{B}) \\ \beta/\varepsilon & -\gamma/\varepsilon \end{pmatrix}. \quad (4.35)$$

On pourrait calculer analytiquement les valeurs propres de cette matrice, mais puisqu'il serait difficile d'obtenir analytiquement les coordonnées du point fixe d'intérêt, on ne pourrait pas les y évaluer. Cependant, on peut facilement obtenir une estimation numérique de la position du point fixe en trouvant numériquement les zéros du champ vectoriel donné par l'équation différentielle (4.17) pour les paramètres choisis. On peut alors calculer numériquement les valeurs propres de la matrice jacobienne (4.35). On obtient les valeurs propres illustrées à la figure 4.5. D'après ces résultats, les parties réelles des valeurs propres passent de négatives à positives en un certain ε près de 1/2, donc il se produit donc une bifurcation de Hopf en ce point.

Pour mieux la visualiser, il est possible d'obtenir une estimation numérique du diagramme de bifurcation. Pour ce faire, on a trouvé numériquement la position du point fixe, puis pour différentes valeurs de ε on a ajouté une petite perturbation à cette valeur, et on a intégré le système (4.17) sur un long intervalle de temps. Lorsque le point fixe est stable, la solution converge vers le point fixe, mais lorsqu'il devient instable, un cycle limite se forme autour du point fixe et la solution converge plutôt vers ce cycle. Ceci permet d'obtenir la position du cycle limite pour différentes valeurs de ε . En répétant cette procédure pour de nombreuses valeurs de ε , on obtient une estimation du diagramme de bifurcation, donnée à la figure 4.6. On y voit bien la forme de la bifurcation et le cycle limite qui est créé alors que la valeur de ε augmente. D'après le diagramme, cette bifurcation de Hopf est supercritique.

On conclut cette sous-section en remarquant que, même si l'existence du cycle limite présenté dans cet exemple est une illustration très intéressante des différences qui existent entre les systèmes (4.17) et (4.19), il n'est pas clair si ces différences peuvent être associées à des propriétés du processus stochastique sous-jacent aux deux modèles. En effet, des trajectoires du processus stochastique ont été obtenues à partir de simulations numériques en choisissant les paramètres d'après (4.31) et l'état initial d'après (4.32). Cependant, dans tous

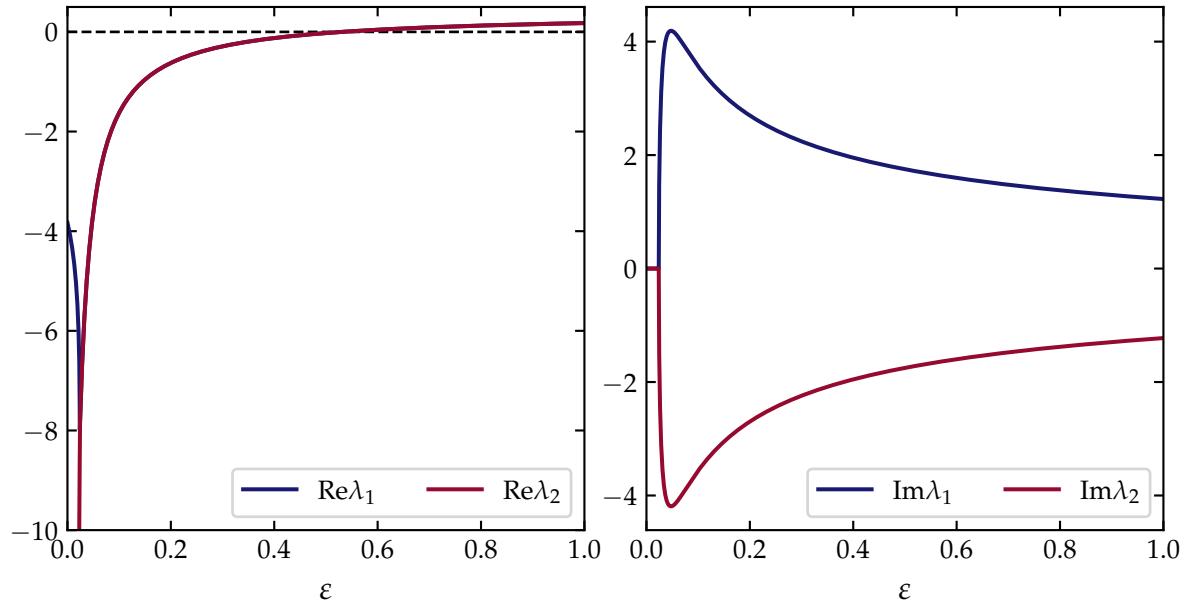


FIGURE 4.5 – Valeurs propres de la matrice jacobienne (4.35), pour les paramètres (4.31), évaluée au point fixe vers lequel la solution converge dans le cas $\varepsilon = 0$. Lorsque les parties imaginaires deviennent non nulles, les courbes des parties réelles sont superposées.

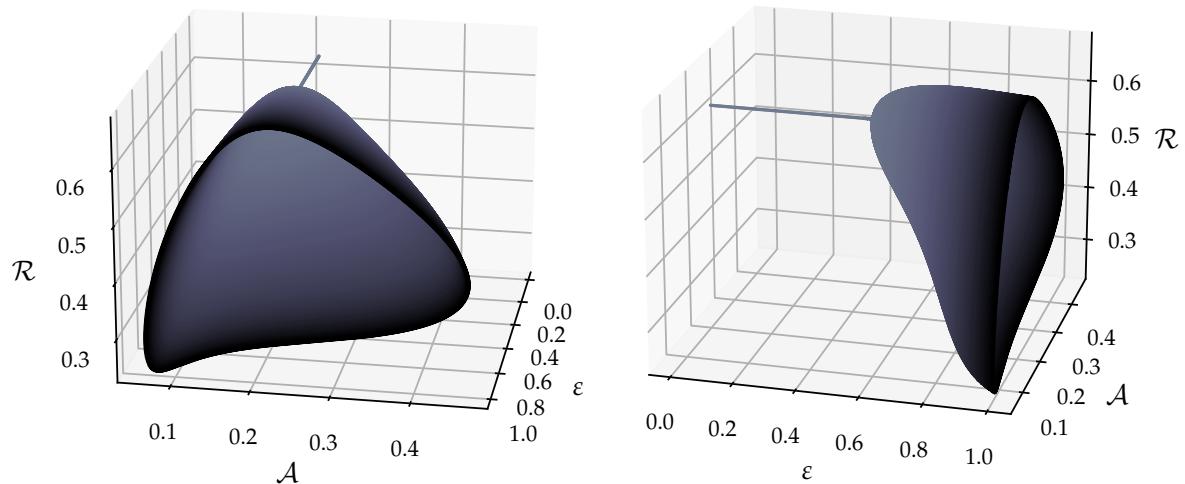


FIGURE 4.6 – Deux angles de vue du diagramme de bifurcation du système (4.34) selon le paramètre ε , en gardant fixés les paramètres (4.31). La couleur de la surface correspond à une fonction de ε qui est utilisée pour que l'on distingue mieux la surface.

les cas observés, les trajectoires du processus stochastique ne correspondaient aux solutions d'aucun des deux systèmes. À la place, on observe à chaque fois une décroissance des fractions de neurones actifs, puis réfractaires, jusqu'à zéro, et aucun neurone ne se réactive. Dans ce cas, on conclut donc qu'aucun des deux modèles n'est une bonne approximation du comportement du processus stochastique.

§ 4.3.3. Étude d'un exemple à deux populations

Des cas semblables au précédent peuvent aussi être vus avec davantage de populations. Par exemple, on considère ici un cas où le graphe est divisé en deux populations notées E et I , qui réfèrent respectivement à des populations de neurones excitateurs et de neurones inhibiteurs. On suppose que les seuils suivent des lois logistiques dans chacune des deux populations, de sorte que les fonctions de répartition F_{θ_E} et F_{θ_I} des deux populations aient toutes les deux la forme (4.31a). On fixe également les paramètres

$$\alpha_E = 10 \text{ ms}^{-1}, \quad \beta_E = 4/5 \text{ ms}^{-1}, \quad \gamma_E = 4 \text{ ms}^{-1}, \quad \theta_E = 0, \quad s_{\theta_E} = 2/5, \quad Q_E = 0, \quad (4.36a)$$

$$\alpha_I = 9 \text{ ms}^{-1}, \quad \beta_I = 1 \text{ ms}^{-1}, \quad \gamma_I = 1 \text{ ms}^{-1}, \quad \theta_I = 3, \quad s_{\theta_I} = 2/5, \quad Q_I = 0 \quad (4.36b)$$

et la matrice de connexion

$$c = \begin{pmatrix} 8 & -12 \\ 9 & -2 \end{pmatrix}. \quad (4.36c)$$

On choisit ensuite l'état initial

$$\mathcal{Y}_0 = (\mathcal{A}_{E0}, \mathcal{A}_{I0}, \mathcal{R}_{E0}, \mathcal{R}_{I0}) := (2/5, 2/5, 2/25, 2/5). \quad (4.37)$$

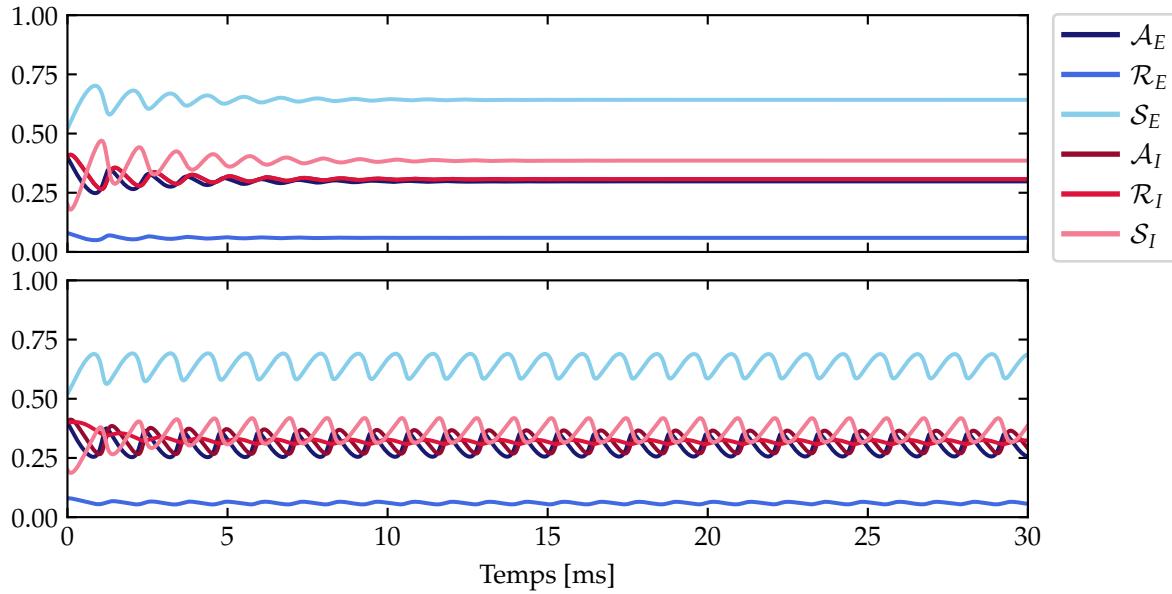
De la même façon que dans l'exemple précédent,

$$\mathcal{R}_{E0} = \frac{2}{25} = \frac{4}{20} \cdot \frac{2}{5} = \frac{\beta_E}{\gamma_E} \mathcal{A}_{E0} \quad \text{et} \quad \mathcal{R}_{I0} = \frac{2}{5} = \frac{\beta_I}{\gamma_I} \mathcal{A}_{I0}, \quad (4.38)$$

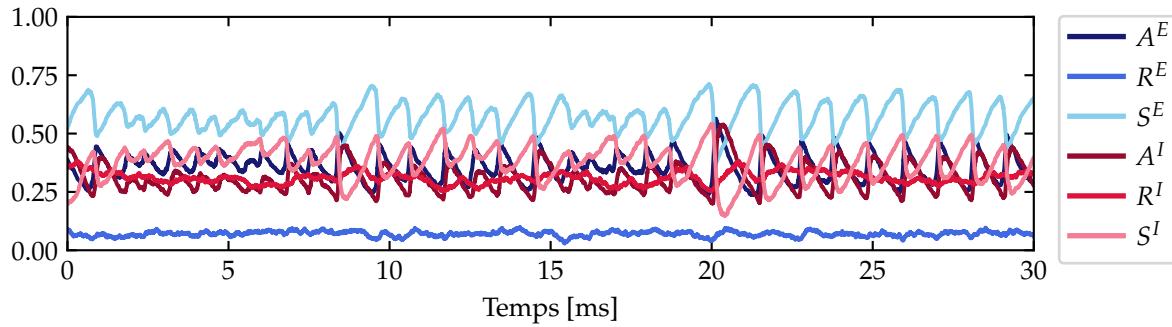
de sorte que $\mathcal{Y}_0 \in \mathcal{D}_2^{\text{WC}}$.

Les solutions des systèmes (4.17) et (4.19) à partir de la condition initiale (4.37) ont été obtenues numériquement, et sont présentées à la figure 4.7a. D'après la figure, le modèle de champ moyen prévoit que la solution converge vers un cycle limite, alors que l'autre modèle prévoit plutôt que la solution converge vers un point fixe stable.

Il est également intéressant de comparer ces solutions à des trajectoires du processus stochastique sous-jacent aux deux modèles, qui peuvent être obtenues par des simulations numériques. Un résultat possible est présenté à la figure 4.7b. On voit alors bien que la trajectoire oscille, et qu'elle semble plus près de la solution du modèle de champ moyen (4.17) que celle du modèle où les fractions réfractaires sont fixées en fonction des fractions actives.



(a) En haut, solution de (4.19) pour les paramètres (4.36) et la condition initiale (4.37). En bas, solution de (4.17) pour les mêmes paramètres et la même condition initiale.



(b) Trajectoire possible du processus de Markov construit à la section 4.2 pour les paramètres (4.36) et l'état initial donné par (4.37), pour un réseau de 1000 neurones par population. Voir l'annexe B.2 pour les détails sur la méthodologie employée pour les simulations.

FIGURE 4.7 – Comparaison entre les solutions des systèmes (4.17) et (4.19) et une trajectoire du processus de Markov construit à la section 4.2 pour les paramètres (4.36) et l'état initial (4.37).

Comme dans le dernier cas, pour mieux comprendre ce qui se produit on combine les deux modèles à partir du système

$$\dot{\mathcal{A}}_J = -\beta_J \mathcal{A}_J + \alpha_J (1 - \mathcal{A}_J - \mathcal{R}_J) F_{\theta_J}(\mathcal{B}_J), \quad (4.39a)$$

$$\varepsilon \dot{\mathcal{R}}_J = -\gamma_J \mathcal{R}_J + \beta_J \mathcal{A}_J. \quad (4.39b)$$

Le cas $\varepsilon = 0$ correspond alors au système (4.19), et le cas $\varepsilon = 1$ correspond au système (4.17).

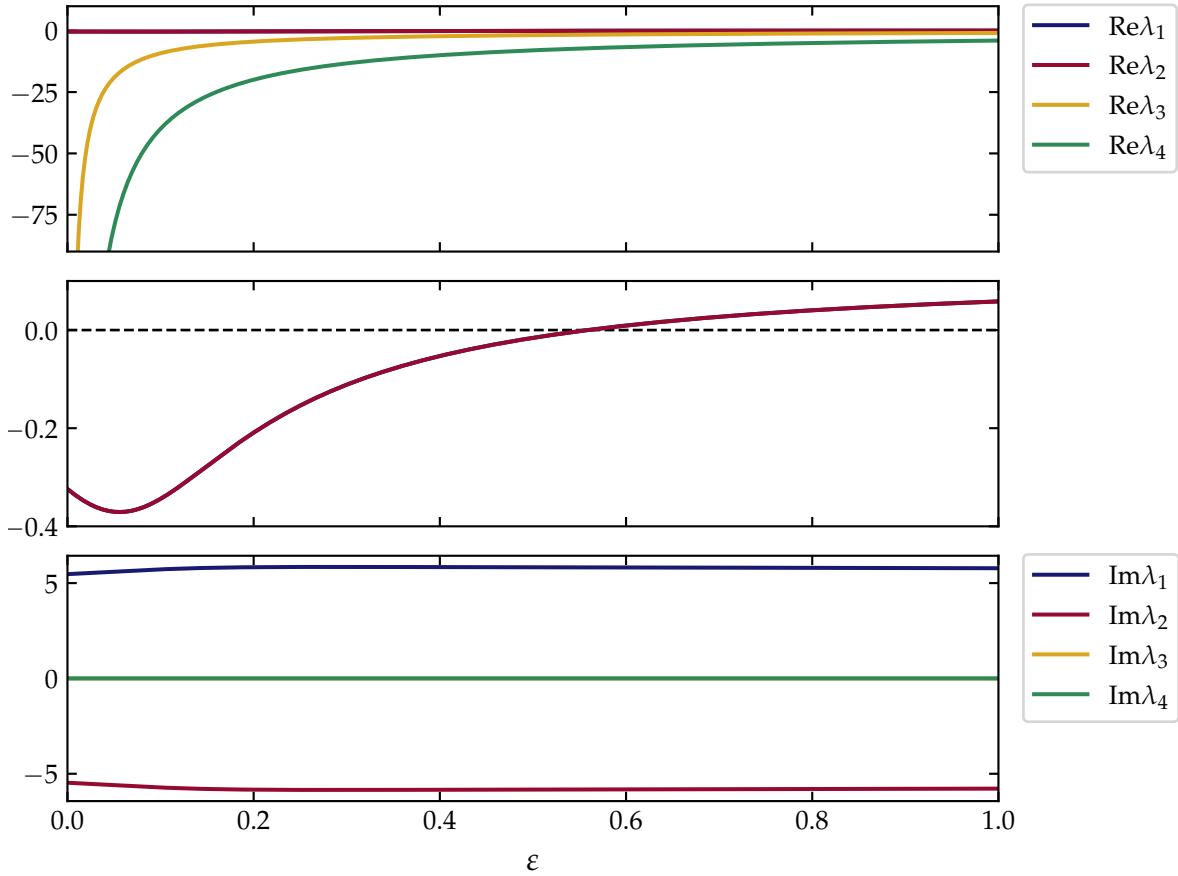


FIGURE 4.8 – Valeurs propres de la matrice jacobienne (4.40) avec les paramètres (4.36) évaluées au point fixe vers lequel la solution converge dans le cas $\varepsilon = 0$. Les parties réelles sont illustrées en haut et au milieu, et les parties imaginaires sont illustrées en bas. Les parties réelles de λ_1 et λ_2 sont égales, et λ_3 et λ_4 sont toutes les deux réelles.

La matrice jacobienne du système combiné (4.39) est donnée par

$$\begin{pmatrix} \xi_E & \alpha_E F'_{\theta_E}(\mathcal{B}_E) c_{EI} \mathcal{S}_E & -\alpha_E F_{\theta_E}(\mathcal{B}_E) & 0 \\ \alpha_I F'_{\theta_I}(\mathcal{B}_I) c_{IE} \mathcal{S}_I & \xi_I & 0 & -\alpha_I F_{\theta_I}(\mathcal{B}_I) \\ \beta_E/\varepsilon & 0 & -\gamma_E/\varepsilon & 0 \\ 0 & \beta_I/\varepsilon & 0 & -\gamma_I/\varepsilon \end{pmatrix}, \quad (4.40)$$

où l'on a défini

$$\xi_J := \beta_J - \alpha_J F_{\theta_J}(\mathcal{B}_J) + \alpha_J F'_{\theta_J}(\mathcal{B}_J) c_{JJ} \mathcal{S}_J \quad (4.41)$$

pour $J \in \{E, I\}$ afin d'éviter les débordements.

Les valeurs propres de cette matrice peuvent être calculées numériquement en fonction de ε de la même façon que dans le dernier cas. Les résultats sont donnés à la figure 4.8. D'après ces résultats, la partie réelle d'une paire de valeurs propres conjuguées change de signe près de la valeur $\varepsilon = 0.55$, donc il se produit une bifurcation de Hopf en ce point.

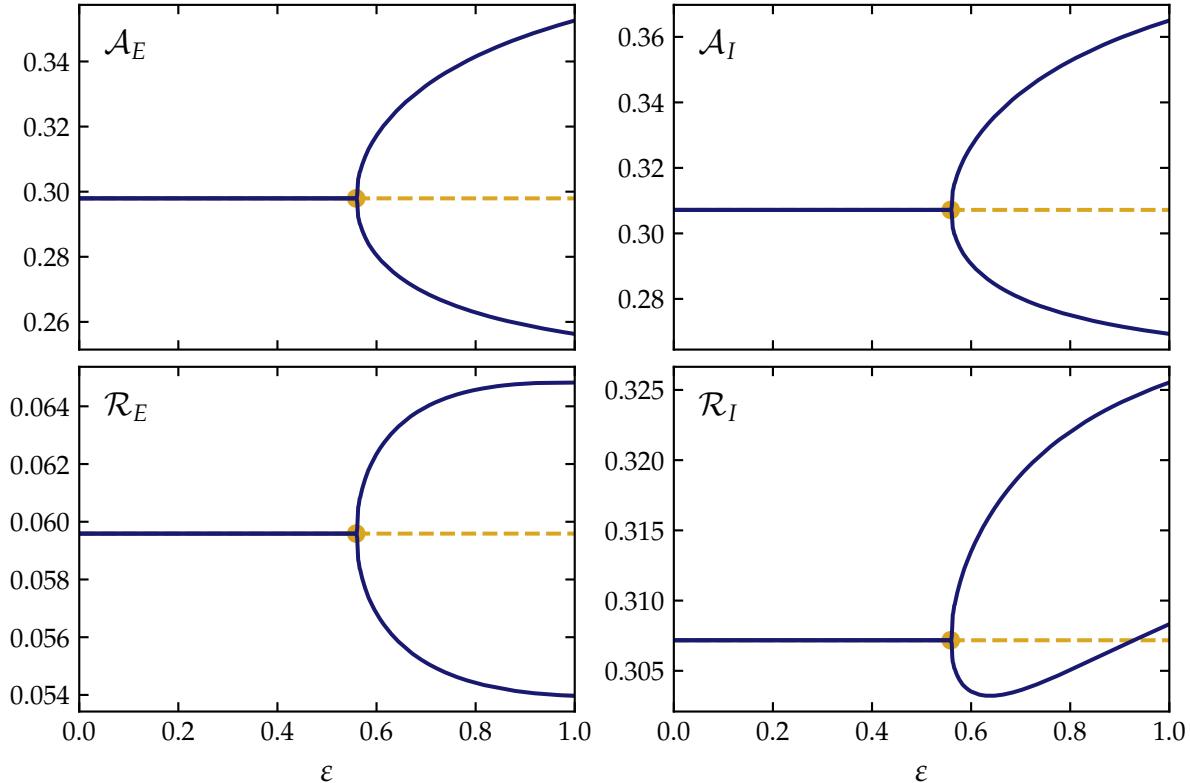


FIGURE 4.9 – Diagrammes de bifurcation du système (4.39) selon ε , pour chacune des quatre composantes. Dans chaque cas, le point jaune illustre le point de bifurcation, et la ligne pointillée indique la valeur de la composante au point fixe. La ligne pleine indique le maximum et le minimum de la solution sur un petit intervalle à la fin de l'intervalle d'intégration.

On peut mieux visualiser la bifurcation en traçant une estimation du diagramme de bifurcation. Cependant, comme le système dynamique a ici quatre dimensions, on ne peut pas procéder de la même façon que dans le dernier cas pour tracer le diagramme. On a donc plutôt calculé, pour chaque valeur de ε après la bifurcation, les valeurs maximale et minimale de chaque composante sur le cycle. Les résultats obtenus sont donnés à la figure 4.9. D'après la forme de ces projections du diagramme de bifurcation, la bifurcation de Hopf est supercritique.

§ 4.3.4. Conclusion

Après la comparaison du système de champ moyen (4.17) et du système (4.19) où les composantes associées aux fractions réfractaires des populations sont fixées d'avance à leurs solutions d'équilibre en fonction des fractions actives, on a vu que ces deux modèles présentent certainement plusieurs similarités importantes. En particulier, on a montré à la proposition 4.1 que leurs points fixes sont les mêmes. Cependant, à la lumière du corollaire de la proposition 4.2, on a vu que la dynamique du système (4.17) ne peut pas être réduite d'emblée à celle du système (4.19).

En effet, les exemples étudiés aux sous-sections 4.3.2 et 4.3.3 montrent que les points fixes des deux systèmes peuvent ne pas avoir la même stabilité même s'ils sont aux mêmes points. Pour étudier la dynamique complète du système (4.17), il est donc nécessaire de tenir compte des fractions de populations associées à l'état réfractaire.

§ 4.4. Approximation de champ moyen de deuxième ordre

À la section 4.2, afin de décrire la dynamique d'un réseau de neurones biologiques d'un point de vue macroscopique, on a proposé le système (4.16), donné par

$$\dot{A}_J(t) = -\beta_J A_J(t) + \alpha_J \mathbb{E}[S_t^J F_{\theta_J}(B_t^J)], \quad (4.16a)$$

$$\dot{\mathcal{R}}_J(t) = -\gamma_J \mathcal{R}_J(t) + \beta_J A_J(t), \quad (4.16b)$$

$$\dot{S}_J(t) = -\alpha_J \mathbb{E}[S_t^J F_{\theta_J}(B_t^J)] + \gamma_J \mathcal{R}_J(t). \quad (4.16c)$$

Comme le modèle construit à la section 4.2 est stochastique, il est naturel de vouloir trouver une solution au problème de fermeture de moments qui irait au-delà de l'approximation de champ moyen. On pourrait alors considérer des corrélations non nulles entre les états de différents neurones du réseau. Il est important de remarquer que, du moins à la connaissance de l'auteur de ce document, ce problème de fermeture de moments n'a jamais été étudié auparavant. Comme discuté plus tôt, les auteurs qui tentent d'inclure des corrélations dans les modèles de réseaux de neurones biologiques semblables au modèle de Wilson–Cowan utilisent plutôt d'autres méthodes, comme des intégrales fonctionnelles ou des équations de Langevin.

Pour attaquer ce problème, il sera plus approprié d'utiliser les moments centrés. Avec les taux de transition moyens donnés en (4.15), le système (2.62) prend la forme

$$\dot{C}_{AA}^{JK}(t) = -(\beta_J + \beta_K) C_{AA}^{JK}(t) + \alpha_K \text{Cov}[A_t^J, S_t^K F_{\theta_K}(B_t^K)] + \alpha_J \text{Cov}[S_t^J F_{\theta_J}(B_t^J), A_t^K], \quad (4.42a)$$

$$\dot{C}_{RR}^{JK}(t) = -(\gamma_J + \gamma_K) C_{RR}^{JK}(t) + \beta_K C_{AR}^{KJ}(t) + \beta_J C_{AR}^{JK}(t), \quad (4.42b)$$

$$\dot{C}_{AR}^{JK}(t) = -(\beta_J + \gamma_K) C_{AR}^{JK}(t) + \beta_K C_{AA}^{JK}(t) + \alpha_J \text{Cov}[S_t^J F_{\theta_J}(B_t^J), R_t^K], \quad (4.42c)$$

où l'on reprend la notation $C_{UV}^{JK}(t) := \text{Cov}[U_t^J, V_t^K]$ où U et V remplacent A , R ou S . Il est important de rappeler que ce système est approximatif au même titre que le système (4.16), en considérant des paramètres homogènes sur la division en populations \mathcal{P} .

Pour trouver une solution au problème de fermeture de moments, on cherche alors à approcher les membres de droite des équations (4.16) et (4.42). Il est intéressant de remarquer d'abord que les approximations utilisées dans le contexte du modèle SIR, présentées à la section 3.3, ne s'appliquent pas bien à ce problème. En effet, ces approximations permettent seulement d'approximer un moment d'ordre 3 en fonction de moments d'ordre 1 et 2. Dans le contexte du modèle SIR, ces approximations sont tout à fait appropriées,

puisque les équations (3.12) et (3.18) qui donnent les dérivées des moments d'ordre 1 et 2 ne contiennent que des moments d'ordre 1, 2 et 3. Or, les équations (4.16) et (4.42) ont une forme très différente en ce qu'elles incluent les fonctions de répartition des seuils. Si l'on suppose que les seuils suivent des lois de probabilité unimodales et symétriques, ces fonctions de répartition ont des formes sigmoïdales. Dans ce cas, les espérances de produits qui impliquent ces fonctions ne dépendent pas que des moments d'ordre 1, 2 et 3, mais dépendent en fait de moments d'ordre arbitrairement élevé.

On propose dans cette section une première solution à ce problème de fermeture de moments, qui peut être vue comme une approximation de champ moyen de deuxième ordre. Il est important de noter que, même si l'on utilise la même expression pour désigner cette approximation, elle est distincte de l'approximation étudiée à la section 3.3.2. Dans la première sous-section, on présente d'abord l'approximation et le système résultant pour un nombre quelconque de populations. Dans la seconde sous-section, on se concentre sur le cas particulier à une population, pour lequel plusieurs résultats analytiques et numériques sont présentés. On présente ensuite un exemple à deux populations à la troisième sous-section. Finalement, on discute les résultats analytiques et numériques qu'on a obtenus et on conclut sur la capacité de ce modèle à représenter la dynamique macroscopique d'un réseau de neurones biologiques.

§ 4.4.1. La fermeture de moments

L'idée de l'approximation présentée ici est assez simple, et peut être vue comme une généralisation de l'approximation de champ moyen qui permette d'inclure les seconds moments. L'approximation de champ moyen consiste à négliger les variances des variables aléatoires d'intérêt, ce qui mène à négliger tous les moments centrés d'ordre supérieur. De la même façon, on néglige ici les moments centrés d'ordre 3 et plus des variables aléatoires d'intérêt.

Pour utiliser cette idée afin d'approximer les membres de droite des équations (4.16) et (4.42), on doit donc les réécrire afin qu'ils s'expriment en fonction des moments centrés des composantes de l'état macroscopique. Pour ce faire, le plus simple est d'utiliser les développements de Taylor des fonctions de répartition des seuils des différentes populations du graphe. Pour simplifier, on suppose que ces fonctions sont toutes analytiques.

Afin d'approximer (4.16a) et (4.16c), on cherche d'abord à approximer les espérances de la forme $\mathbb{E}[S_t^J F_{\theta_J}(B_t^J)]$. En exprimant F_{θ_J} d'après sa série de Taylor autour de l'espérance $\mathcal{B}_J(t)$ de B_t^J ,

$$\mathbb{E}[S_t^J F_{\theta_J}(B_t^J)] = \mathbb{E}\left[S_t^J \sum_{k=0}^{\infty} \frac{1}{k!} F_{\theta_J}^{(k)}(\mathcal{B}_J(t)) (B_t^J - \mathcal{B}_J(t))^k\right]. \quad (4.43)$$

En supposant que F_{θ_J} est telle que l'espérance peut être distribuée dans la série, on voit que

$$\begin{aligned}\mathbb{E}[S_t^J F_{\theta_J}(B_t^J)] &= \sum_{k=0}^{\infty} \frac{1}{k!} F_{\theta_J}^{(k)}(\mathcal{B}_J(t)) \mathbb{E}[(S_t^J - \mathcal{S}_J(t))(B_t^J - \mathcal{B}_J(t))^k] \\ &\quad + \sum_{k=0}^{\infty} \frac{1}{k!} F_{\theta_J}^{(k)}(\mathcal{B}_J(t)) \mathcal{S}_J(t) \mathbb{E}[(B_t^J - \mathcal{B}_J(t))^k].\end{aligned}\tag{4.44}$$

On remarque alors que le terme $k = 0$ de la première série est identiquement nul puisque $\mathbb{E}[S_t^J - \mathcal{S}_J(t)] = 0$, de même que le terme $k = 1$ de la seconde série. Ensuite, comme on veut négliger les moments centrés d'ordre supérieur à deux, on néglige les termes où $k \geq 2$ dans la première série, et ceux où $k \geq 3$ dans la seconde. Ainsi,

$$\begin{aligned}\mathbb{E}[S_t^J F_{\theta_J}(B_t^J)] &\approx F'_{\theta_J}(\mathcal{B}_J(t)) \mathbb{E}[(S_t^J - \mathcal{S}_J(t))(B_t^J - \mathcal{B}_J(t))] \\ &\quad + \mathcal{S}_J(t) F_{\theta_J}(\mathcal{B}_J(t)) + \frac{1}{2} \mathcal{S}_J(t) F''_{\theta_J}(\mathcal{B}_J(t)) \mathbb{E}[(B_t^J - \mathcal{B}_J(t))^2]\end{aligned}\tag{4.45}$$

$$= \mathcal{S}_J(t) F_{\theta_J}(\mathcal{B}_J(t)) + F'_{\theta_J}(\mathcal{B}_J(t)) C_{SB}^{JJ}(t) + \frac{1}{2} \mathcal{S}_J(t) F''_{\theta_J}(\mathcal{B}_J(t)) C_{BB}^{JJ}(t).\tag{4.46}$$

Il est important de remarquer ici que cette approximation ne dépend que des composantes A_t^J et R_t^J de l'état macroscopique, puisque S_t^J et B_t^J en dépendent linéairement. En effet,

$$C_{SB}^{JJ}(t) = \text{Cov}[S_t^J, B_t^J] = \text{Cov}\left[1 - A_t^J - R_t^J, \sum_{K \in \mathcal{P}} c_{JK} A_t^K + Q_J\right]\tag{4.47}$$

$$= - \sum_{K \in \mathcal{P}} c_{JK} \left(\text{Cov}[A_t^J, A_t^K] + \text{Cov}[R_t^J, A_t^K] \right) = - \sum_{K \in \mathcal{P}} c_{JK} (C_{AA}^{JK}(t) + C_{AR}^{JK}(t)).\tag{4.48}$$

De la même façon,

$$C_{BB}^{JJ}(t) = \text{Cov}[B_t^J, B_t^J] = \text{Cov}\left[\sum_{K \in \mathcal{P}} c_{JK} A_t^K + Q_J, \sum_{L \in \mathcal{P}} c_{JL} A_t^L + Q_J\right] = \sum_{K, L \in \mathcal{P}} c_{JK} c_{JL} C_{AA}^{KL}(t).\tag{4.49}$$

L'approximation de $\mathbb{E}[S_t^J F_{\theta_J}(B_t^J)]$ ne dépend ainsi que des premiers et seconds moments des composantes de l'état macroscopique.

Les mêmes idées peuvent être utilisées pour approximer les covariances qui impliquent des fonctions de répartition dans (4.42). D'abord,

$$\text{Cov}[A_t^J, S_t^K F_{\theta_K}(B_t^K)] = \mathbb{E}\left[(A_t^J - \mathcal{A}_J(t)) \left(S_t^K F_{\theta_K}(B_t^K) - \mathbb{E}[S_t^K F_{\theta_K}(B_t^K)]\right)\right]\tag{4.50}$$

$$\begin{aligned}&= \mathbb{E}\left[(A_t^J - \mathcal{A}_J(t)) (S_t^K - \mathcal{S}_K(t)) F_{\theta_K}(B_t^K)\right] \\ &\quad + \mathcal{S}_K(t) \mathbb{E}\left[(A_t^J - \mathcal{A}_J(t)) F_{\theta_K}(B_t^K)\right].\end{aligned}\tag{4.51}$$

Dans la première espérance, seul le terme constant donne un moment centré d'ordre inférieur à trois lorsqu'on développe F_{θ_K} en série de Taylor. Dans la deuxième, le terme constant s'annule, et seul le terme linéaire donne un moment centré d'ordre inférieur à trois. On obtient ainsi l'approximation

$$\text{Cov}[A_t^J, S_t^K F_{\theta_K}(B_t^K)] \approx F_{\theta_K}(\mathcal{B}_K(t)) C_{AS}^{JK}(t) + \mathcal{S}_K(t) F'_{\theta_K}(\mathcal{B}_K(t)) C_{AB}^{JK}(t).\tag{4.52}$$

En remplaçant A_t^J par R_t^J , la même démarche mène à l'approximation

$$\text{Cov}[R_t^J, S_t^K F_{\theta_K}(B_t^K)] \approx F_{\theta_K}(\mathcal{B}_K(t)) C_{RS}^{JK}(t) + \mathcal{S}_K(t) F'_{\theta_K}(\mathcal{B}_K(t)) C_{RB}^{JK}(t). \quad (4.53)$$

En utilisant les approximations (4.46) pour $\mathbb{E}[S_t^J F_{\theta_J}(B_t^J)]$, (4.52) pour $\text{Cov}[A_t^J, S_t^K F_{\theta_K}(B_t^K)]$ et (4.53) pour $\text{Cov}[R_t^J, S_t^K F_{\theta_K}(B_t^K)]$ dans le système (4.16)–(4.42), on obtient le système

$$\dot{\mathcal{A}}_J = -\beta_J \mathcal{A}_J + \alpha_J \mathcal{S}_J F_{\theta_J}(\mathcal{B}_J) + \alpha_J F'_{\theta_J}(\mathcal{B}_J) C_{SB}^{JJ} + \frac{\alpha_J}{2} \mathcal{S}_J F''_{\theta_J}(\mathcal{B}_J) C_{BB}^{JJ}, \quad (4.54a)$$

$$\dot{\mathcal{R}}_J = -\gamma_J \mathcal{R}_J + \beta_J \mathcal{A}_J, \quad (4.54b)$$

$$\begin{aligned} \dot{C}_{AA}^{JK} = & -(\beta_J + \beta_K) C_{AA}^{JK} + \alpha_K F_{\theta_K}(\mathcal{B}_K) C_{AS}^{JK} + \alpha_K \mathcal{S}_K F'_{\theta_K}(\mathcal{B}_K) C_{AB}^{JK} \\ & + \alpha_J F_{\theta_J}(\mathcal{B}_J) C_{AS}^{KJ} + \alpha_J \mathcal{S}_J F'_{\theta_J}(\mathcal{B}_J) C_{AB}^{KJ}, \end{aligned} \quad (4.54c)$$

$$\dot{C}_{RR}^{JK} = -(\gamma_J + \gamma_K) C_{RR}^{JK} + \beta_K C_{AR}^{KJ} + \beta_J C_{AR}^{JK}, \quad (4.54d)$$

$$\dot{C}_{AR}^{JK} = -(\beta_J + \gamma_K) C_{AR}^{JK} + \beta_K C_{AA}^{JK} + \alpha_J F_{\theta_J}(\mathcal{B}_J) C_{RS}^{KJ} + \alpha_J \mathcal{S}_J F'_{\theta_J}(\mathcal{B}_J) C_{RB}^{KJ}. \quad (4.54e)$$

À partir de la relation $S_t^J = 1 - A_t^J - R_t^J$ et de l'expression $B_t^J = \sum_{L \in \mathcal{P}} c_{JL} A_t^L + Q_J$ de l'activité entrante, on peut réécrire toutes les covariances qui se trouvent dans ces équations en fonction des covariances entre des fractions actives et réfractaires de populations du graphe. Ainsi, le système (4.54) ne dépend que des espérances et des seconds moments centrés des composantes A_t^J et R_t^J de l'état macroscopique.

Il est clair que si l'on néglige toutes les covariances dans le système (4.54), on retombe précisément sur le système (4.17) obtenu de l'approximation de champ moyen. On peut donc voir le système (4.54) comme une extension du modèle de champ moyen.

Il est facile de déterminer la dimension du système (4.54). D'abord, en notant n le nombre de populations que comporte le graphe, on voit qu'il y a n équations de la forme (4.54a) et n équations de la forme (4.54b) qui tiennent respectivement compte des composantes \mathcal{A}_J et \mathcal{R}_J . Ensuite, on rappelle que pour toutes populations J et K du graphe, $C_{AA}^{JK} = C_{AA}^{KJ}$. Ainsi, il y a $\frac{n(n+1)}{2}$ équations de la forme (4.54c) pour les composantes C_{AA}^{JK} . De la même façon, il y a $\frac{n(n+1)}{2}$ équations de la forme (4.54d) pour les composantes C_{RR}^{JK} . Finalement, il y a aussi n^2 équations de la forme (4.54e) pour les composantes C_{AR}^{JK} . Ainsi, le système (4.54) est de dimension $2n + 2\frac{n(n+1)}{2} + n^2 = n(2n + 3)$.

On peut d'emblée remarquer que, comme on suppose que les fonctions de répartition F_{θ_J} sont analytiques, le champ vectoriel associé au système (4.54) est de classe \mathcal{C}^1 . Ainsi, d'après le théorème 1.2, ce système admet une solution unique pour toute condition initiale dans $\mathbb{R}^{n(2n+3)}$. Pour l'étude analytique du système, on s'attaque au cas le plus simple, où le graphe n'a qu'une population.

§ 4.4.2. Étude du cas à une population

On étudie ici le système dynamique (4.54) pour le cas d'une seule population. Dans ce cas, le système a cinq dimensions, et il est simple de l'écrire en utilisant seulement des covariances entre les fractions active et réfractaire de la population. On obtient alors le système

$$\dot{\mathcal{A}} = -\beta\mathcal{A} + \alpha SF_\theta(\mathcal{B}) - \alpha F'_\theta(\mathcal{B})c(C_{AA} + C_{AR}) + \frac{\alpha}{2}SF''_\theta(\mathcal{B})c^2C_{AA}, \quad (4.55a)$$

$$\dot{\mathcal{R}} = -\gamma\mathcal{R} + \beta\mathcal{A}, \quad (4.55b)$$

$$\dot{C}_{AA} = -2\beta C_{AA} - 2\alpha F_\theta(\mathcal{B})(C_{AA} + C_{AR}) + 2\alpha SF'_\theta(\mathcal{B})cC_{AA}, \quad (4.55c)$$

$$\dot{C}_{RR} = -2\gamma C_{RR} + 2\beta C_{AR}, \quad (4.55d)$$

$$\dot{C}_{AR} = -(\beta + \gamma)C_{AR} + \beta C_{AA} - \alpha F_\theta(\mathcal{B})(C_{RR} + C_{AR}) + \alpha SF'_\theta(\mathcal{B})cC_{AR}, \quad (4.55e)$$

où l'on a laissé tombé les indices qui réfèreraient à l'unique population. Ce système est étudié particulièrement dans l'idée de le comparer au système (4.17), obtenu dans l'approximation de champ moyen. On présente ici quelques résultats sur l'analyse de ce système dynamique, notamment sur certains domaines invariants par son flot ainsi que sur ses points fixes. On discute ensuite quelques résultats obtenus numériquement et qui visaient à trouver des comportements distincts de ceux prévus par le système (4.17).

§ 4.4.2.1. Domaine physiologique

Bien que le système dynamique décrit par (4.55) soit défini dans \mathbb{R}^5 , les variables dynamiques n'ont un sens que dans un domaine restreint de \mathbb{R}^5 , à la manière de ce qu'on a décrit à la section 4.3.1 lorsqu'on a défini le domaine \mathcal{D}_n en (4.23). De la même façon qu'à cette section, on veut qu'en tout temps $\mathcal{A}, \mathcal{R}, \mathcal{S} \in [0, 1]$, c'est-à-dire que $(\mathcal{A}, \mathcal{R}) \in \mathcal{D}_1$.

D'autres conditions peuvent toutefois être obtenues sur les variances des fractions active et réfractaire ainsi que sur leur covariance. D'abord, puisque les variables aléatoires A_t et R_t sont bornées entre 0 et 1 pour tout $t \geq 0$, l'inégalité de Bhatia–Davis impose que

$$C_{AA} \leq \mathcal{A}(1 - \mathcal{A}) \quad \text{et que} \quad C_{RR} \leq \mathcal{R}(1 - \mathcal{R}). \quad (4.56)$$

En utilisant le même raisonnement pour la fraction sensible, on voit que

$$C_{SS} = C_{AA} + 2C_{AR} + C_{RR} \leq \mathcal{S}(1 - \mathcal{S}) \iff C_{AR} \leq \frac{1}{2}(\mathcal{S}(1 - \mathcal{S}) - C_{AA} - C_{RR}). \quad (4.57)$$

La covariance entre les fractions active et réfractaire peut ensuite être bornée à partir de l'inégalité de Cauchy–Schwarz. En effet, celle-ci impose que $|C_{AR}| \leq \sqrt{C_{AA}C_{RR}}$. L'inégalité s'applique aussi aux autres covariances qui ne sont pas des variances, mais ces cas sont redondants. En effet, puisque $C_{AR} = -C_{AA} - C_{AS}$ et que $C_{RR} = C_{AA} + 2C_{AS} + C_{SS}$, l'inégalité de Cauchy–Schwarz pour C_{AR} montre que

$$(C_{AA} + C_{AS})^2 = C_{AR}^2 \leq C_{AA}C_{RR} = C_{AA}(C_{AA} + 2C_{AS} + C_{SS}). \quad (4.58)$$

Or, en détaillant les deux côtés, on obtient l'inégalité

$$C_{AA}^2 + 2C_{AA}C_{AS} + C_{AS}^2 \leq C_{AA}^2 + 2C_{AA}C_{AS} + C_{AA}C_{SS} \quad (4.59)$$

qui, une fois simplifiée, est exactement l'inégalité $C_{AS}^2 \leq C_{AA}C_{SS}$. Une démarche équivalente montre que $C_{AR}^2 \leq C_{AA}C_{RR}$ implique que $C_{RS}^2 \leq C_{RR}C_{SS}$.

D'autres conditions peuvent encore borner la covariance entre les fractions active et réfractaire. D'abord, puisque A_t et R_t sont positives pour tout t , leur produit est aussi positif. Ainsi, $C_{AR} = \mathbb{E}[A_tR_t] - \mathcal{AR} \geq -\mathcal{AR}$. De la même façon,

$$C_{AS} = -C_{AA} - C_{AR} \geq -\mathcal{AS} \iff C_{AR} \leq \mathcal{AS} - C_{AA} \quad (4.60)$$

et

$$C_{RS} = -C_{RR} - C_{AR} \geq -\mathcal{RS} \iff C_{AR} \leq \mathcal{RS} - C_{RR}. \quad (4.61)$$

On note d'ailleurs que les inégalités (4.60) et (4.61) impliquent ensemble l'inégalité (4.57) ; en les sommant, on obtient que

$$C_{AR} = \frac{1}{2}(C_{AR} + C_{AS}) \leq \frac{1}{2}(\mathcal{AS} - C_{AA} + \mathcal{RS} - C_{RR}), \quad (4.62)$$

qui correspond à (4.57), puisque $\mathcal{A} + \mathcal{R} = 1 - \mathcal{S}$.

Il est important de remarquer que les bornes inférieures sur C_{AR} données par l'inégalité de Cauchy–Schwarz et par la relation $C_{AR} \geq -\mathcal{AR}$ sont indépendantes, au sens où elles peuvent l'une et l'autre être atteinte sans que l'autre ne le soit. D'abord, la borne donnée par l'inégalité de Cauchy–Schwarz peut être atteinte sans que $C_{AR} = -\mathcal{AR}$, puisque les variables aléatoires A_t et R_t pourraient être parfaitement anticorrélées sans que l'espérance de leur produit ne soit nulle. Pour le sens inverse, le plus simple est de prendre un exemple.

Il serait possible que

$$(A_t, R_t) = \begin{cases} (0, 0) & \text{avec probabilité } p_S, \\ (1, 0) & \text{avec probabilité } p_A, \\ (0, 1) & \text{avec probabilité } p_R, \end{cases} \quad (4.63)$$

où $p_S + p_A + p_R = 1$. Dans ce cas,

$$\mathbb{E}[A_tR_t] = \sum_{(a,r) \in \{0,1\}^2} ar \mathbb{P}[A_t = a, R_t = r] = 0. \quad (4.64)$$

Cependant, $\mathbb{E}[A_t] = \mathbb{E}[A_t^2] = p_A$ et $\mathbb{E}[R_t] = \mathbb{E}[R_t^2] = p_R$, donc $C_{AA} = p_A(1 - p_A)$ et $C_{RR} = p_R(1 - p_R)$. Il s'ensuit que $C_{AR} = -p_A p_R$. De plus, puisque $1 - p_A > 1 - p_A - p_S = p_R$ et que $1 - p_R > 1 - p_R - p_S = p_A$ si $p_S > 0$, alors

$$C_{AR} = -p_A p_R = -\sqrt{p_A^2 p_R^2} > -\sqrt{p_A(1 - p_A)p_R(1 - p_R)} = -\sqrt{C_{AA}C_{RR}}. \quad (4.65)$$

La borne inférieure $C_{AR} \geq -\mathcal{AR}$ peut donc être atteinte sans que la borne inférieure donnée par l'inégalité de Cauchy–Schwarz ne le soit. Les mêmes raisonnements s'appliquent évidemment aux autres covariances.

En résumé, la covariance C_{AR} doit respecter les conditions

$$C_{AR} \geq -\sqrt{C_{AA}C_{RR}}, \quad C_{AR} \geq -\mathcal{A}\mathcal{R}, \quad (4.66a)$$

$$C_{AR} \leq \sqrt{C_{AA}C_{RR}}, \quad C_{AR} \leq \mathcal{A}\mathcal{S} - C_{AA}, \quad \text{et} \quad C_{AR} \leq \mathcal{R}\mathcal{S} - C_{RR}. \quad (4.66b)$$

En rassemblant les conditions énumérées plus haut, on arrive à la conclusion que les variables dynamiques ont un sens seulement dans le domaine

$$\mathcal{D}_1^+ := \left\{ (\mathcal{A}, \mathcal{R}, C_{AA}, C_{RR}, C_{AR}) \in [0, 1]^4 \times [-1, 1] \mid \begin{array}{l} \mathcal{A} + \mathcal{R} \leq 1, C_{AR} \text{ respecte (4.66),} \\ C_{AA} \leq \mathcal{A}(1 - \mathcal{A}), C_{RR} \leq \mathcal{R}(1 - \mathcal{R}) \end{array} \right\}. \quad (4.67)$$

On appellera parfois cet ensemble le *domaine physiologique*, puisque c'est le domaine où les variables dynamiques ont un sens du point de vue de l'étude du réseau de neurones biologiques.

§ 4.4.2.2. Ensembles invariants

Dans un premier temps, il est intéressant de chercher si certains ensembles sont laissés invariants par le flot du système dynamique (4.55), particulièrement si ces ensembles ont un sens du point de vue biologique.

Par exemple, il serait intéressant d'arriver à montrer que le domaine \mathcal{D}_1^+ défini en (4.67) est invariant. Ceci assurerait que les solutions du système qui débutent dans ce domaine y restent toujours, c'est-à-dire que le système ne prévoit que des solutions correspondant à des trajectoires possibles du processus stochastique sous-jacent. Malheureusement, la non linéarité du système (4.55) et la complexité du domaine (4.67) rendent cette analyse difficile, et aucun résultat en ce sens n'a pu être obtenu.

On peut cependant obtenir quelques résultats moins forts, et tout de même intéressants. Le résultat suivant permet de voir le cas d'une seule population du système (4.17) obtenu de l'approximation de champ moyen comme un sous-système de (4.55).

Proposition 4.3. *Le plan $C_{AA} = C_{RR} = C_{AR} = 0$ dans \mathbb{R}^5 est invariant par le flot du système (4.55). ◆*

Démonstration. Il est clair à partir des expressions des dérivées de C_{AA} , C_{RR} et C_{AR} données en (4.55) qu'en tout point où $C_{AA} = C_{RR} = C_{AR} = 0$, les dérivées \dot{C}_{AA} , \dot{C}_{RR} et \dot{C}_{AR} sont toutes nulles. Ainsi, le champ vectoriel donné par (4.55) est parallèle au plan $C_{AA} = C_{RR} = C_{AR} = 0$ dans \mathbb{R}^5 . Par le théorème 1.4, ce plan est donc invariant par le flot de (4.55). ■

Le résultat suivant permet de voir que certaines bornes sur la covariance entre les fractions active et réfractaire sont toujours respectées par une solution dont la condition initiale les respecte.

Proposition 4.4. *Les ensembles*

$$F_{\pm} := \left\{ (\mathcal{A}, \mathcal{R}, C_{AA}, C_{RR}, C_{AR}) \in \mathbb{R}^5 : C_{AR} = \pm \sqrt{C_{AA}C_{RR}} \neq 0 \right\} \quad (4.68)$$

sont invariants par le flot du système (4.55). \blacklozenge

Démonstration. Soient $g_{\pm}: \mathbb{R}^2 \times (0, \infty)^2 \times \mathbb{R} \rightarrow \mathbb{R}$ les fonctions définies par

$$g_{\pm}(\mathcal{A}, \mathcal{R}, C_{AA}, C_{RR}, C_{AR}) := \sqrt{C_{AA}C_{RR}} \mp C_{AR}. \quad (4.69)$$

Soient x une solution du système (4.54) et t tel que $x(t)$ se trouve dans le domaine de g_{\pm} . Alors

$$\dot{g}_{\pm}(x(t)) = \langle Dg_{\pm}(x(t)), \dot{x}(t) \rangle = \langle Dg_{\pm}(x(t)), f(x(t)) \rangle \quad (4.70)$$

où $\langle \cdot, \cdot \rangle$ note le produit scalaire usuel sur \mathbb{R}^5 et où $f: \mathbb{R}^5 \rightarrow \mathbb{R}^5$ note le champ vectoriel associé au système (4.55). Un calcul direct montre que

$$Dg_{\pm}(\mathcal{A}, \mathcal{R}, C_{AA}, C_{RR}, C_{AR}) = \left(0, 0, \frac{1}{2} \sqrt{\frac{C_{RR}}{C_{AA}}}, \frac{1}{2} \sqrt{\frac{C_{AA}}{C_{RR}}}, \mp 1 \right). \quad (4.71)$$

Ainsi, en notant $x(t) = (\mathcal{A}(t), \mathcal{R}(t), C_{AA}(t), C_{RR}(t), C_{AR}(t))$, on voit que l'on peut écrire un peu plus explicitement

$$\dot{g}_{\pm}(x(t)) = \frac{1}{2} \sqrt{\frac{C_{RR}(t)}{C_{AA}(t)}} \dot{C}_{AA} \Big|_{x(t)} + \frac{1}{2} \sqrt{\frac{C_{AA}(t)}{C_{RR}(t)}} \dot{C}_{RR} \Big|_{x(t)} \mp \dot{C}_{AR} \Big|_{x(t)}, \quad (4.72)$$

où \dot{C}_{AA} , \dot{C}_{RR} et \dot{C}_{AR} sont les fonctions données par le système (4.55). En omettant les dépendances sur t pour alléger la notation, on calcule alors que

$$\begin{aligned} \dot{g}_{\pm}(x) &= \frac{1}{2} \sqrt{\frac{C_{RR}}{C_{AA}}} \left(-2\beta C_{AA} - 2\alpha F_{\theta}(\mathcal{B})(C_{AA} + C_{AR}) + 2\alpha S F'_{\theta}(\mathcal{B})c C_{AA} \right) \\ &\quad + \frac{1}{2} \sqrt{\frac{C_{AA}}{C_{RR}}} \left(-2\gamma C_{RR} + 2\beta C_{AR} \right) \end{aligned} \quad (4.73)$$

$$\begin{aligned} &\pm (\beta + \gamma) C_{AR} \mp \beta C_{AA} \pm \alpha F_{\theta}(\mathcal{B})(C_{RR} + C_{AR}) \mp \alpha S F'_{\theta}(EB)c C_{AR} \\ &= -\beta \sqrt{C_{AA}C_{RR}} - \alpha F_{\theta}(\mathcal{B}) \sqrt{C_{AA}C_{RR}} - \alpha F_{\theta} C_{AR} \sqrt{\frac{C_{RR}}{C_{AA}}} + \alpha S F'_{\theta}(\mathcal{B})c \sqrt{C_{AA}C_{RR}} \\ &\quad - \gamma \sqrt{C_{AA}C_{RR}} + \beta C_{AR} \sqrt{\frac{C_{AA}}{C_{RR}}} \end{aligned} \quad (4.74)$$

$$\begin{aligned} &\pm (\beta + \gamma) C_{AR} \mp \beta C_{AA} \pm \alpha F_{\theta}(\mathcal{B})(C_{RR} + C_{AR}) \mp \alpha S F'_{\theta}(EB)c C_{AR} \\ &= - \left(\beta + \alpha F_{\theta}(\mathcal{B}) + \gamma \pm \beta \sqrt{\frac{C_{AA}}{C_{RR}}} \mp \alpha F_{\theta}(\mathcal{B}) \sqrt{\frac{C_{RR}}{C_{AA}}} - \alpha S F'_{\theta}(\mathcal{B})c \right) g_{\pm}(x). \end{aligned} \quad (4.75)$$

Il est alors clair que les ensembles où $g_{\pm} = 0$ sont invariants par le flot du système (4.55), puisque si $g_{\pm}(x) = 0$, alors $\dot{g}_{\pm}(x) = 0$. Or, $g_{\pm}(\mathcal{A}, \mathcal{R}, C_{AA}, C_{RR}, C_{AR}) = 0$ si et seulement si $C_{AR} = \pm \sqrt{C_{AA}C_{RR}}$. Ainsi, on conclut que les ensembles F_{\pm} sont invariants par le flot du système (4.55). \blacksquare

Ce résultat implique d'ailleurs qu'une solution de l'équation différentielle (4.55) dont la condition initiale se trouve dans le domaine \mathcal{D}_1^+ , qui respecte forcément l'inégalité de Cauchy–Schwarz, respectera toujours cette inégalité. Cependant, il est difficile de s'assurer que les variances — en particulier C_{AA} — n'augmenteront pas en dehors des bornes du domaine \mathcal{D}_1^+ .

§ 4.4.2.3. Points fixes du système dynamique

On étudie maintenant les points fixes de l'équation différentielle (4.55). Le lemme suivant montre qu'il est possible de classer ces points fixes en trois catégories, la première correspondant aux points fixes du système de champ moyen (4.17). Les deux autres catégories sont particulièrement intéressantes, puisque dans ces cas les composantes associées aux covariances pourraient être non nulles. De tels points fixes pourraient donc correspondre à des états stables du système stochastique sous-jacent qui ne peuvent être modélisés par le système de champ moyen.

Lemme 4.5. *On suppose que $(\mathcal{A}^*, \mathcal{R}^*, C_{AA}^*, C_{RR}^*, C_{AR}^*) \in \mathbb{R}^5$ est un point fixe de (4.55), et pour simplifier la notation, on note $F^{(k)} = F_\theta^{(k)}(\mathcal{B}^*)$. Alors $\mathcal{R}^* = \frac{\beta}{\gamma}\mathcal{A}^*$ et $C_{RR}^* = \frac{\beta}{\gamma}C_{AR}^*$. De plus, l'une des conditions suivantes doit être respectée :*

i) $C_{AA}^* = C_{RR}^* = C_{AR}^* = 0$, et \mathcal{A}^* est une solution de l'équation

$$\alpha F_\theta(\mathcal{B}) = \frac{\beta\gamma\mathcal{A}}{\gamma - (\beta + \gamma)\mathcal{A}}. \quad (4.76)$$

ii) $C_{AA}^* = \frac{\alpha F}{\gamma}C_{AR}^*$,

$$C_{AR}^* = \frac{\gamma}{\beta + \gamma} \frac{(\beta + \gamma + \alpha F)(\beta\gamma + \beta\alpha F + \gamma\alpha F) - \beta\gamma\alpha F'c}{(\gamma + \alpha F)(\alpha F'c)^2 - \frac{1}{2}(\beta + \gamma + \alpha F)\alpha^2 FF''c^2}, \quad (4.77)$$

et \mathcal{A}^* est une solution de l'équation $\alpha F'_\theta(\mathcal{B})c(1 - (1 + \frac{\beta}{\gamma})\mathcal{A}) = \beta + \gamma + \alpha F_\theta(\mathcal{B})$.

iii) $C_{AA}^* = \frac{\gamma}{\beta}C_{AR}^*$,

$$C_{AR}^* = \frac{\beta}{\gamma(\beta + \gamma)} \frac{(\beta\gamma + \beta\alpha F + \gamma\alpha F)^2 - \beta\gamma^2\alpha F'c}{(\beta + \gamma)(\alpha F'c)^2 - \frac{1}{2}(\beta\gamma + \beta\alpha F + \gamma\alpha F)\alpha F''c^2}, \quad (4.78)$$

et \mathcal{A}^* est une solution de l'équation $\gamma\alpha F'_\theta(\mathcal{B})c(1 - (1 + \frac{\beta}{\gamma})\mathcal{A}) = \beta\gamma + \beta\alpha F_\theta(\mathcal{B}) + \gamma\alpha F_\theta(\mathcal{B})$. ◆

Démonstration. D'abord, il est clair à partir des expressions (4.55b) et (4.55d) des dérivées de \mathcal{R} et de C_{RR} que $\mathcal{R}^* = \frac{\beta}{\gamma}\mathcal{A}^*$ et que $C_{RR}^* = \frac{\beta}{\gamma}C_{AR}^*$. Ensuite, en évaluant la dérivée (4.55e) de C_{AR} au point fixe et en substituant $C_{RR}^* = \frac{\beta}{\gamma}C_{AR}^*$, on obtient la relation

$$0 = -(\beta + \gamma)C_{AR}^* + \beta C_{AA}^* - \alpha F\left(\frac{\beta}{\gamma} + 1\right)C_{AR}^* + \alpha S^*F'cC_{AR}^* \quad (4.79)$$

$$= \beta C_{AA}^* - \left(\beta + \gamma + \alpha F + \frac{\beta\alpha F}{\gamma} - \alpha F'cS^*\right)C_{AR}^*, \quad (4.80)$$

où l'on utilise la notation $F^{(k)} = F_\theta^{(k)}(\mathcal{B}^*)$ introduite dans l'énoncé du lemme. On conclut

alors que

$$C_{AA}^* = \frac{1}{\beta\gamma} (\beta\gamma + \gamma^2 + \gamma\alpha F + \beta\alpha F - \gamma\alpha F'c\mathcal{S}^*) C_{AR}^*. \quad (4.81)$$

De plus, en évaluant la dérivée (4.55c) de C_{AA} au point fixe et en y substituant ce résultat, on obtient que

$$0 = -2\beta C_{AA}^* - 2\alpha F(C_{AA}^* + C_{AR}^*) + 2\alpha \mathcal{S}^* F'c C_{AA}^* \quad (4.82)$$

$$= -2\alpha F C_{AR}^* - \frac{2}{\beta\gamma} (\beta + \alpha F - \alpha F'c\mathcal{S}^*) (\beta\gamma + \gamma^2 + \gamma\alpha F + \beta\alpha F - \gamma\alpha F'c\mathcal{S}^*) C_{AR}^* \quad (4.83)$$

$$= -\frac{2}{\beta\gamma} ((\beta\gamma + \beta\alpha F + \gamma\alpha F - \gamma\alpha F'c\mathcal{S}^*) (\beta + \alpha F - \alpha F'c\mathcal{S}^*) \\ + \gamma(\beta\alpha F + \beta\gamma + \gamma\alpha F - \gamma\alpha F'c\mathcal{S}^*)) C_{AR}^*. \quad (4.84)$$

Ainsi, on trouve l'équation

$$0 = \frac{2}{\beta\gamma} (\beta + \gamma + \alpha F - \alpha F'c\mathcal{S}^*) (\beta\gamma + \beta\alpha F + \gamma\alpha F - \gamma\alpha F'c\mathcal{S}^*) C_{AR}^*. \quad (4.85)$$

On a alors trois possibilités. D'abord, il se peut que $C_{AR}^* = 0$, auquel cas les relations trouvées plus tôt montrent que $C_{AA}^* = C_{RR}^* = 0$. Dans ce cas, on obtient la valeur de \mathcal{A}^* à partir de la dérivée de \mathcal{A} évaluée au point fixe. En évaluant (4.55a) en un point $(\mathcal{A}, \frac{\beta}{\gamma}\mathcal{A}, 0, 0, 0)$, on obtient que

$$\dot{\mathcal{A}} = -\beta\mathcal{A} + \alpha \left(1 - \left(1 + \frac{\beta}{\gamma} \right) \mathcal{A} \right) F_\theta(\mathcal{B}) = -\beta\mathcal{A} + \alpha \left(\frac{\gamma - (\beta + \gamma)\mathcal{A}}{\gamma} \right) F_\theta(\mathcal{B}). \quad (4.86)$$

Ainsi, $(\mathcal{A}^*, \frac{\beta}{\gamma}\mathcal{A}^*, 0, 0, 0)$ est un point fixe du système (4.55) si \mathcal{A}^* est une solution de (4.76).

Ensuite, l'équation (4.85) peut aussi être respectée si \mathcal{A}^* est une solution de l'équation

$$\alpha F'_\theta(\mathcal{B})c \left(1 - \left(1 + \frac{\beta}{\gamma} \right) \mathcal{A} \right) = \beta + \gamma + \alpha F_\theta(\mathcal{B}), \quad (4.87)$$

c'est-à-dire que $\alpha F'c\mathcal{S}^* = \beta + \gamma + \alpha F$. Dans ce cas, on obtient alors de (4.81) que $C_{AA}^* = \frac{\alpha F}{\gamma} C_{AR}^*$. On peut ensuite trouver une expression pour C_{AR}^* en fonction de \mathcal{A}^* à partir de la dérivée (4.55a) de \mathcal{A} évaluée au point fixe. En y substituant $C_{AA}^* = \frac{\alpha F}{\gamma} C_{AR}^*$, on trouve que

$$0 = -\beta\mathcal{A}^* + \alpha \mathcal{S}^* F - \alpha F'c \left(\frac{\alpha F}{\gamma} + 1 \right) C_{AR}^* + \frac{\alpha}{2} \mathcal{S}^* F''c^2 \frac{\alpha F}{\gamma} C_{AR}^*. \quad (4.88)$$

En utilisant alors la relation $\mathcal{S}^* = 1 - \frac{\beta+\gamma}{\gamma}\mathcal{A}^*$, équivalente à $\mathcal{A}^* = \frac{\gamma}{\beta+\gamma}(1 - \mathcal{S}^*)$, on voit que

$$0 = \frac{(\beta\gamma + \beta\alpha F + \gamma\alpha F)\mathcal{S}^* - \beta\gamma}{\beta + \gamma} - \frac{1}{\gamma} \left(\alpha F'c(\gamma + \alpha F) - \frac{1}{2} \alpha^2 F F'' c^2 \mathcal{S}^* \right) C_{AR}^*. \quad (4.89)$$

On isole alors C_{AR}^* pour trouver que

$$C_{AR}^* = \frac{\gamma}{\beta + \gamma} \frac{(\beta\gamma + \beta\alpha F + \gamma\alpha F)\alpha F'c\mathcal{S}^* - \beta\gamma\alpha F'c}{(\gamma + \alpha F)(\alpha F'c)^2 - \frac{1}{2}\alpha^2 F F'' c^2 \alpha F'c \mathcal{S}^*}. \quad (4.90)$$

On retrouve finalement l'expression annoncée pour C_{AR}^* dans le cas *i)* en substituant $\alpha F' c \mathcal{S}^* = \beta + \gamma + \alpha F$ dans l'expression précédente.

Finalement, l'équation (4.85) peut être respectée dans un troisième cas : lorsque \mathcal{A}^* est une solution de l'équation

$$\gamma \alpha F'_\theta(\mathcal{B}) c \left(1 - \left(1 + \frac{\beta}{\gamma} \right) \mathcal{A}^* \right) = \beta \gamma + \beta \alpha F_\theta(\mathcal{B}) + \gamma \alpha F_\theta(\mathcal{B}), \quad (4.91)$$

c'est-à-dire que $\gamma \alpha F' c \mathcal{S}^* = \beta \gamma + \beta \alpha F + \gamma \alpha F$. Dans ce cas, on obtient de (4.81) que $C_{AA}^* = \frac{\gamma}{\beta} C_{AR}^*$. Comme plus tôt, on trouve une expression pour C_{AR}^* à partir de la dérivée (4.55a) de \mathcal{A} évaluée au point fixe. En y substituant $C_{AA}^* = \frac{\gamma}{\beta} C_{AR}^*$, on trouve que

$$0 = -\beta \mathcal{A}^* + \alpha \mathcal{S}^* F - \alpha F' c \left(\frac{\gamma}{\beta} + 1 \right) C_{AR}^* + \frac{\alpha}{2} \mathcal{S}^* F'' c^2 \frac{\gamma}{\beta} C_{AR}^* \quad (4.92)$$

$$= \frac{(\beta \gamma + \beta \alpha F + \gamma \alpha F) \mathcal{S}^* - \beta \gamma}{\beta + \gamma} - \frac{1}{\beta} \left(\alpha F' c (\beta + \gamma) - \frac{1}{2} \gamma \alpha F'' c^2 \mathcal{S}^* \right) C_{AR}^*, \quad (4.93)$$

où l'on a utilisé la relation $\mathcal{A}^* = \frac{\gamma}{\beta + \gamma} (1 - \mathcal{S}^*)$ comme plus tôt. On isole alors C_{AR}^* pour trouver que

$$C_{AR}^* = \frac{\beta}{\beta + \gamma} \frac{(\beta \gamma + \beta \alpha F + \gamma \alpha F) \gamma \alpha F' c \mathcal{S}^* - \beta \gamma^2 \alpha F' c}{\gamma (\beta + \gamma) (\alpha F' c)^2 - \frac{1}{2} \gamma^2 \alpha F'' c^2 \alpha F' c \mathcal{S}^*}, \quad (4.94)$$

et l'on retrouve l'expression annoncée en substituant $\gamma \alpha F' c \mathcal{S}^* = \beta \gamma + \beta \alpha F + \gamma \alpha F$. ■

Les points fixes de la forme *i)* sont les plus faciles à étudier. Le résultat suivant donne un critère pour leur existence, et même pour leur stabilité.

Proposition 4.6. *Tous les points fixes du système (4.55) de la forme *i)* décrite au lemme 4.5 se trouvent dans le domaine \mathcal{D}_1^+ , et il en existe toujours au moins un. De plus, un point fixe de cette forme est stable si sa coordonnée \mathcal{A}^* respecte la condition*

$$\beta + \gamma + \alpha F_\theta(\mathcal{B}^*) - \alpha F'_\theta(\mathcal{B}^*) c \left(1 - \left(1 + \frac{\beta}{\gamma} \right) \mathcal{A}^* \right) > \max \left\{ 0, \gamma - \frac{\beta \alpha F_\theta(\mathcal{B}^*)}{\gamma} \right\}, \quad (4.95)$$

et il est instable si l'inégalité stricte inverse est respectée. ◆

Remarque 1. Pour une population de neurones inhibiteurs, c'est-à-dire si $c < 0$, la condition de stabilité est forcément respectée. En effet, dans ce cas le membre de gauche de (4.95) est directement positif, et la condition est forcément vraie puisque F_θ est positive. On peut interpréter ceci biologiquement en ce que si la valeur de \mathcal{A} est légèrement supérieure à sa valeur au point fixe, la population de neurones s'inhibe elle-même davantage, ce qui a tendance à ralentir l'activation d'autres neurones, donc à faire diminuer la valeur de \mathcal{A} vers sa valeur au point fixe. À l'inverse, si la valeur de \mathcal{A} un légèrement inférieure à sa valeur au point fixe, la population de neurones s'inhibe moins, ce qui facilite l'activation des neurones et fait augmenter la valeur de \mathcal{A} .

Remarque 2. Si $\beta > \alpha c \sup F'_\theta$, la condition de stabilité est forcément respectée. Pour le voir, on suppose que $c \geq 0$ puisque le cas $c < 0$ est couvert par la remarque précédente. Alors le membre de gauche de (4.95) est toujours supérieur à $\beta + \gamma - \alpha F'_\theta(\mathcal{B}^*)c$, puisque F_θ et F'_θ sont positives et que $\mathcal{S}^* = 1 - \frac{\beta+\gamma}{\gamma} \mathcal{A}^*$ est borné entre 0 et 1. Ainsi, la condition de stabilité est respectée si $\beta > \alpha F'_\theta(\mathcal{B}^*)c$. Pour comprendre l'interprétation biologique de cette condition dans le cas où $c \geq 0$, on remarque que si $\Delta \mathcal{A}$ est petit, alors par une approximation de Taylor, $\alpha F_\theta(c(\mathcal{A}^* + \Delta \mathcal{A}) + Q) \approx \alpha F_\theta(\mathcal{B}^*) + \alpha F'_\theta(\mathcal{B}^*)c\Delta \mathcal{A}$. Une petite perturbation $\Delta \mathcal{A}$ de la fraction de neurones actifs a donc pour effet de modifier le taux d'activation d'environ $\alpha F'_\theta(\mathcal{B}^*)c\Delta \mathcal{A}$, qui a toujours le signe de $\Delta \mathcal{A}$. Ainsi, lorsque $\beta > \alpha c \sup F'_\theta$, les neurones se désactivent trop rapidement pour qu'une modification de la fraction de neurones actifs puisse avoir un effet durable sur la dynamique du réseau.

Démonstration. D'après le lemme 4.5, un point fixe de la forme i) doit avoir les coordonnées $(\mathcal{A}^*, \frac{\beta}{\gamma} \mathcal{A}^*, 0, 0, 0)$ pour un certain $\mathcal{A}^* \in \mathbb{R}$. L'existence d'un tel point fixe dans le domaine \mathcal{D}_1^+ correspond à l'existence d'une solution à l'équation (4.76) telle que $\mathcal{A}^*, \mathcal{R}^*, \mathcal{S}^* \in [0, 1]$. Puisque $\mathcal{S}^* = 1 - \mathcal{A}^* - \mathcal{R}^* = 1 - \frac{\beta+\gamma}{\gamma} \mathcal{A}^*$, alors $0 \leq \mathcal{S}^* \leq 1$ si $0 \leq \mathcal{A}^* \leq \frac{\gamma}{\beta+\gamma} < 1$, auquel cas $0 \leq \mathcal{R}^* = \frac{\beta}{\gamma} \mathcal{A}^* \leq \frac{\beta}{\beta+\gamma} < 1$. Ainsi, il existe un point fixe de la forme i) dans \mathcal{D}_1^+ s'il existe une solution à l'équation (4.76) dans l'intervalle $[0, \frac{\gamma}{\beta+\gamma}]$.

On veut donc montrer que la fonction $g: [0, \frac{\gamma}{\beta+\gamma}] \rightarrow \mathbb{R}$ donnée par

$$g(\mathcal{A}) := \alpha F_\theta(c\mathcal{A} + Q) - \frac{\beta\gamma\mathcal{A}}{\gamma - (\beta + \gamma)\mathcal{A}} \quad (4.96)$$

a toujours un zéro. On remarque d'abord que g est continue, puisque F_θ est continue. De plus,

$$\frac{\beta\gamma\mathcal{A}}{\gamma - (\beta + \gamma)\mathcal{A}} \rightarrow \infty \quad \text{lorsque} \quad \mathcal{A} \rightarrow \frac{\gamma}{\beta + \gamma}, \quad (4.97)$$

alors que cette fonction vaut 0 lorsque \mathcal{A} vaut 0. Par ailleurs, F_θ est bornée entre 0 et 1, puisqu'elle est une fonction de répartition. On conclut donc que $g(0) \geq 0$ et que $g(\mathcal{A}) \rightarrow -\infty$ lorsque $\mathcal{A} \rightarrow \frac{\gamma}{\beta+\gamma}$, ce qui implique que $g(\mathcal{A}) < 0$ pour un certain $\mathcal{A} \in [0, \frac{\gamma}{\beta+\gamma}]$. Par le théorème des valeurs intermédiaires, g doit donc avoir un zéro dans $[0, \frac{\gamma}{\beta+\gamma}]$. Le système (4.55) admet donc toujours au moins un point fixe de forme i) dans le domaine \mathcal{D}_1^+ .

On remarque également que le signe de $\beta\gamma\mathcal{A}$ suit le signe de \mathcal{A} , alors que le signe de $\gamma - (\beta + \gamma)\mathcal{A}$ suit le signe de $\mathcal{A} - \frac{\gamma}{\beta+\gamma}$. Ainsi, le membre de droite de (4.76) est non négatif seulement si $\mathcal{A} \in [0, \frac{\gamma}{\beta+\gamma}]$. Comme F_θ est non négative, ceci montre que tous les zéros de g se trouvent dans l'intervalle $\mathcal{A} \in [0, \frac{\gamma}{\beta+\gamma}]$. Tous les points fixes de forme i) se trouvent donc dans le domaine \mathcal{D}_1^+ .

La stabilité d'un tel point fixe peut ensuite être étudiée à partir des valeurs propres de la

matrice jacobienne. Une fois évaluée au point fixe, celle-ci est donnée par

$$\mathcal{J}^* := \begin{pmatrix} -\beta - \xi + \xi' \mathcal{S}^* & -\xi & -\xi' + \xi'' & 0 & -\xi' \\ \beta & -\gamma & 0 & 0 & 0 \\ 0 & 0 & -2\beta - 2\xi + 2\xi' \mathcal{S}^* & 0 & -2\xi \\ 0 & 0 & 0 & -2\gamma & 2\beta \\ 0 & 0 & \beta & -\xi & -\beta - \gamma - \xi + \xi' \mathcal{S}^* \end{pmatrix}, \quad (4.98)$$

où l'on a introduit $\xi := \alpha F_\theta(\mathcal{B}^*)$, $\xi' := \alpha F'_\theta(\mathcal{B}^*)c$ et $\xi'' := \frac{\alpha}{2} F''_\theta(\mathcal{B}^*)c^2$ afin d'alléger la notation pour le calcul. Comme \mathcal{J}^* est triangulaire par blocs, on peut séparer le calcul des valeurs propres ; on a en effet que $\text{spec } \mathcal{J}^* = \text{spec } \mathcal{J}_1 \cup \text{spec } \mathcal{J}_2$, où

$$\mathcal{J}_1 := \begin{pmatrix} -\beta - \xi + \xi' \mathcal{S}^* & -\xi \\ \beta & -\gamma \end{pmatrix} \quad (4.99)$$

et

$$\mathcal{J}_2 := \begin{pmatrix} -2\beta - 2\xi + 2\xi' \mathcal{S}^* & 0 & -2\xi \\ 0 & -2\gamma & 2\beta \\ \beta & -\xi & -\beta - \gamma - \xi + \xi' \mathcal{S}^* \end{pmatrix}. \quad (4.100)$$

On commence par obtenir les valeurs propres de \mathcal{J}_1 . On calcule donc le déterminant

$$\det(\mathcal{J}_1 - \lambda I) = (-\beta - \xi + \xi' \mathcal{S}^* - \lambda)(-\gamma - \lambda) + \beta \xi \quad (4.101)$$

$$= \lambda^2 + (\beta + \gamma + \xi - \xi' \mathcal{S}^*)\lambda + \beta\gamma + \beta\xi + \gamma\xi - \gamma\xi' \mathcal{S}^*. \quad (4.102)$$

Ainsi, $\det(\mathcal{J}_1 - \lambda I) = 0$ si et seulement si $\lambda = \lambda_{\pm}$, où

$$\lambda_{\pm} := -\frac{\beta + \gamma + \xi - \xi' \mathcal{S}^*}{2} \pm \sqrt{\left(\frac{\beta + \gamma + \xi - \xi' \mathcal{S}^*}{2}\right)^2 - \beta\gamma - \beta\xi - \gamma\xi + \gamma\xi' \mathcal{S}^*}. \quad (4.103)$$

On obtient ensuite les valeurs propres de \mathcal{J}_2 . De même, on calcule le déterminant

$$\det(\mathcal{J}_2 - \lambda I) = (-2\beta - 2\xi + 2\xi' \mathcal{S}^* - \lambda)(-2\gamma - \lambda)(-\beta - \gamma - \xi + \xi' \mathcal{S}^* - \lambda) + 2\beta\xi(-2\gamma - \lambda) + 2\beta\xi(-2\beta - 2\xi + 2\xi' \mathcal{S}^* - \lambda) \quad (4.104)$$

$$= -(\beta + \gamma + \xi - \xi' \mathcal{S}^* + \lambda)((2\beta + 2\xi - 2\xi' \mathcal{S}^* + \lambda)(2\gamma + \lambda) + 4\beta\xi) \quad (4.105)$$

$$= -4(\beta + \gamma + \xi - \xi' \mathcal{S}^* + \lambda) \det(\mathcal{J}_1 - \frac{\lambda}{2} I). \quad (4.106)$$

Ainsi, $\det(\mathcal{J}_2 - \lambda I) = 0$ si et seulement si $\lambda = 2\lambda_{\pm}$ ou $\lambda = \lambda_0 := -\beta - \gamma - \xi + \xi' \mathcal{S}^*$.

On trouve donc que

$$\text{spec } \mathcal{J}^* = \{\lambda_0, \lambda_+, \lambda_-, 2\lambda_+, 2\lambda_-\}. \quad (4.107)$$

En termes des variables initiales, ces valeurs propres sont données par

$$\lambda_0 = -\beta - \gamma - \alpha F_\theta(\mathcal{B}^*) + \alpha F'_\theta(\mathcal{B}^*)c\mathcal{S}^* \quad (4.108a)$$

et

$$\begin{aligned}\lambda_{\pm} &= \frac{\lambda_0}{2} \pm \sqrt{\frac{\lambda_0^2}{4} - \beta\gamma - \beta\alpha F_\theta(\mathcal{B}^*) - \gamma\alpha F_\theta(\mathcal{B}^*) + \gamma\alpha F'_\theta(\mathcal{B})cES^*} \\ &= \frac{\lambda_0}{2} \left(1 \pm \sqrt{1 - 4 \frac{\beta\gamma + \beta\alpha F_\theta(\mathcal{B}^*) + \gamma\alpha F_\theta(\mathcal{B}^*) - \gamma\alpha F'_\theta(\mathcal{B}^*)cS^*}{(\beta + \gamma + \alpha F_\theta(\mathcal{B}^*) - \alpha F'_\theta(\mathcal{B}^*)cS^*)^2}} \right).\end{aligned}\quad (4.108b)$$

On sait que le point fixe est stable si les parties réelles de toutes les valeurs propres de \mathcal{J}^* sont négatives. Comme λ_0 est forcément réelle, on veut donc que $\lambda_0 < 0$. Cependant, les λ_{\pm} peuvent avoir une partie imaginaire non nulle. Si c'est le cas, alors leur partie réelle vaut $\lambda_0/2$, de sorte que si $\lambda_0 < 0$, alors $\text{Re } \lambda_{\pm} < 0$. Si les λ_{\pm} sont réelles, alors on veut simplement que l'argument de la racine reste inférieur à 1, puisque dans ce cas $\lambda_{\pm} < 0$ si $\lambda_0 < 0$, et dans le cas contraire λ_+ et λ_- sont forcément de signes opposés.

Pour que les parties réelles de toutes les valeurs propres de \mathcal{J}^* soient négatives, il est donc nécessaire et suffisant que

$$\lambda_0 < 0 \quad \text{et que} \quad \beta\gamma + \beta\alpha F_\theta(\mathcal{B}^*) + \gamma\alpha F_\theta(\mathcal{B}^*) - \gamma\alpha F'_\theta(\mathcal{B}^*)cS^* > 0. \quad (4.109)$$

On note que

$$\beta\gamma + \beta\alpha F_\theta(\mathcal{B}^*) + \gamma\alpha F_\theta(\mathcal{B}^*) - \gamma\alpha F'_\theta(\mathcal{B}^*)cS^* = -\gamma\lambda_0 + \beta\alpha F_\theta(\mathcal{B}^*) - \gamma^2, \quad (4.110)$$

de sorte que le second critère de (4.109) est équivalent à $\lambda_0 < \frac{\gamma^2 - \beta\alpha F_\theta(\mathcal{B}^*)}{\gamma}$. Ainsi, le point fixe est stable si et seulement si

$$-\lambda_0 > \max \left\{ 0, \frac{\gamma^2 - \beta\alpha F_\theta(\mathcal{B}^*)}{\gamma} \right\}, \quad (4.111)$$

ce qui correspond au critère annoncé une fois l'expression de λ_0 détaillée. ■

Il est plus difficile d'assurer l'existence de points fixes des deux autres formes décrites au lemme 4.5. On donne toutefois une condition qui permet de les écarter de ce domaine.

Proposition 4.7. Si $\beta > \alpha c \sup F'_\theta$, alors tous les points fixes du système (4.55) dans \mathcal{D}_1^+ sont de la forme i) donnée au lemme 4.5. ◆

Démonstration. On a vu en (4.85) que les coordonnées $(\mathcal{A}^*, \mathcal{R}^*, \mathcal{C}_{AA}^*, \mathcal{C}_{RR}^*, \mathcal{C}_{AR}^*)$ d'un point fixe doivent toujours respecter la relation

$$0 = \frac{2}{\beta\gamma} (\beta + \gamma + \alpha F - \alpha F' c S^*) (\beta\gamma + \beta\alpha F + \gamma\alpha F - \gamma\alpha F' c S^*) \mathcal{C}_{AR}^*, \quad (4.112)$$

où l'on reprend la notation $F^{(k)} = F_\theta^{(k)}(\mathcal{B}^*)$ utilisée plus tôt. On suppose que ce point fixe se trouve dans le domaine \mathcal{D}_1^+ , ce qui implique que $0 \leq S^* \leq 1$. On rappelle que F_θ est bornée entre 0 et 1 et que F'_θ est positive. Ainsi, si $c < 0$, les deux parenthèses dans (4.112)

sont strictement positives. Si $c \geq 0$, alors les mêmes propriétés montrent que

$$\beta + \gamma + \alpha F - \alpha F' c \mathcal{S}^* \geq \beta + \gamma - \alpha c \sup F'_\theta \quad (4.113)$$

et que

$$\beta \gamma + \beta \alpha F + \gamma \alpha F - \gamma \alpha F' c \mathcal{S}^* \geq \beta \gamma - \gamma \alpha c \sup F'_\theta. \quad (4.114)$$

Comme $\beta > \alpha c \sup F'_\theta$, ces deux quantités sont strictement positives. Dans tous les cas, les deux parenthèses dans (4.112) sont strictement positives, de sorte que cette condition ne puisse être respectée que si $C_{AR}^* = 0$ ce qui, comme dans la preuve du lemme 4.5, mène à la conclusion que le point fixe est de forme *i*). ■

Il est beaucoup plus difficile d'obtenir une condition générale qui assurerait l'existence de points fixes de formes *ii*) et *iii*) dans le domaine \mathcal{D}_1^+ . Cependant, ces points fixes sont ceux qui présentent le plus d'intérêt, puisqu'ils permettraient de prédire des états stables d'un réseau de neurones biologiques qui ne pourraient pas être prédits par les modèles de champ moyen.

Pour tenter d'étudier de tels points fixes, on peut en chercher dans des contextes restreints, avec quelques suppositions additionnelles. En effet, on rappelle que des points fixes de formes *ii*) et *iii*) existent lorsque la valeur \mathcal{A}^* au point fixe est respectivement telle que

$$\mathcal{S}^* = \frac{\beta + \gamma + \alpha F_\theta(\mathcal{B}^*)}{\alpha F'_\theta(\mathcal{B}^*)c} \quad \text{ou} \quad \mathcal{S}^* = \frac{\beta \gamma + \beta \alpha F_\theta(\mathcal{B}^*) + \gamma \alpha F_\theta(\mathcal{B}^*)}{\gamma \alpha F'_\theta(\mathcal{B}^*)c}. \quad (4.115)$$

Dans chacun de ces deux cas, pour que \mathcal{S}^* soit entre 0 et 1, il est impératif que la valeur de F'_θ soit assez grande. Si les seuils suivent une loi de probabilité unimodale et symétrique, alors leur densité F'_θ sera plus grande près de leur moyenne θ . On pourrait donc chercher des points fixes pour une activité entrante externe

$$Q = \theta - c \mathcal{A}^* = \theta - \frac{c \gamma}{\beta + \gamma} (1 - \mathcal{S}^*), \quad (4.116)$$

de sorte que $\mathcal{B}^* = c \mathcal{A}^* + Q = \theta$, et que le point fixe se trouve exactement au point d'inflexion de la fonction de répartition. Dans ce cas, $F_\theta(\mathcal{B}^*) = 1/2$ et $F'_\theta(\mathcal{B}^*) = \max F'_\theta$, et les valeurs de \mathcal{S}^* deviennent

$$\mathcal{S}^* = \frac{\alpha + 2\beta + 2\gamma}{2\alpha F'_\theta(\mathcal{B}^*)c} \quad \text{et} \quad \mathcal{S}^* = \frac{\alpha \beta + \alpha \gamma + 2\beta \gamma}{2\gamma \alpha F'_\theta(\mathcal{B}^*)c}. \quad (4.117)$$

Les valeurs prévues pour \mathcal{S}^* en (4.117) ne dépendent que des paramètres. Elles permettent donc toujours de trouver des points fixes de formes *ii*) et *iii*), qui peuvent se trouver dans le domaine \mathcal{D}_1^+ comme ils peuvent ne pas s'y trouver.

Il est possible, à partir de ces valeurs, d'évaluer la matrice jacobienne du système à l'un ou l'autre de ces points fixes afin d'en étudier la stabilité. Étant donnée la taille des matrices, le calcul explicite des valeurs propres n'a pas permis de mener à un critère utile sur la

stabilité des points fixes. Plutôt que de tenter de travailler les valeurs propres elles-mêmes pour trouver des critères, on a donc plutôt tenté d'utiliser les propriétés de la matrice.

Par exemple, un résultat de Wolkowicz et Styan [102] donne des conditions pour que toutes les parties réelles des valeurs propres d'une matrice soient négatives. Malheureusement, avec les deux matrices jacobienes étudiées ici, ces résultats mènent à des conditions suffisantes à la stabilité qui ne peuvent jamais être satisfaites, du moins pour des points fixes qui se trouveraient dans le domaine physiologique \mathcal{D}_1^+ . Wolkowicz et Styan ont proposé d'autres critères dans [103], mais ceux-ci dépendent des normes de fonctions de la matrice d'intérêt, ce qui mène ici à des calculs très lourds, desquels aucun résultat n'a pu être obtenu. On mentionne également que les diagonales des deux matrices jacobienes comportent des entrées positives, de sorte qu'il n'est pas possible d'utiliser le théorème des cercles de Gershgorin [36] (voir aussi le théorème 6.1.1 dans [51] pour un énoncé du théorème en anglais) afin d'obtenir un critère qui montrerait que les points fixes sont stables.

Le seul résultat qui a été obtenu en lien avec la stabilité des points fixes de formes *ii)* ou *iii)* est un critère simple d'instabilité pour un point fixe de forme *iii)* dans le domaine \mathcal{D}_1^+ , dans le cas où ce point fixe est déterminé par (4.117) avec les suppositions qui ont mené à cette expression. Dans ce cas, on peut en effet montrer que la trace de la matrice jacobienne évaluée au point fixe est donnée par

$$2\frac{\alpha\beta}{\gamma} - 4\gamma + \alpha F_\theta'''(\theta)c^3 \frac{\gamma}{2\beta} \mathcal{S}^* \mathbf{C}_{AR}^*. \quad (4.118)$$

Si cette trace est strictement positive, alors la partie réelle d'au moins une valeur propre doit être strictement positive, auquel cas le point fixe doit être instable. Si le point fixe se trouve dans le domaine \mathcal{D}_1^+ , alors $\mathcal{S}^* \leq 1$ et $\mathbf{C}_{AR}^* \leq 1/4$. De plus, comme on suppose que les seuils suivent une loi de probabilité unimodale et symétrique, alors F_θ'' doit être décroissante en θ par la proposition 1.8, de sorte que $F_\theta'''(\theta) \leq 0$. Pour que la trace (4.118) soit strictement positive, il est donc suffisant que

$$2\frac{\alpha\beta}{\gamma} - 4\gamma + \alpha F_\theta'''(\theta)c^3 \frac{\gamma}{2\beta} \frac{1}{4} > 0, \quad (4.119)$$

c'est-à-dire que

$$16\alpha\beta^2 > 32\beta\gamma^2 - \alpha\gamma F_\theta'''(\theta)c^3. \quad (4.120)$$

§ 4.4.2.4. Quelques résultats numériques

Pour avoir une meilleure intuition sur le système (4.55), on a effectué des intégrations numériques des équations différentielles. Ainsi, on a pu explorer les solutions possibles du système pour différentes valeurs des paramètres.

Pour ce faire, on a d'abord obtenu de nombreuses solutions du système en choisissant des paramètres aléatoirement, et en fixant des conditions initiales aléatoirement dans le

domaine \mathcal{D}_1^+ . De ces essais, on a vu que dans la grande majorité des cas, les variances décroissent toutes rapidement vers zéro, et les solutions stables prévues par le modèle (4.55) sont les mêmes que celles prévues par le modèle issu de l'approximation de champ moyen. Dans certains cas des solutions différentes ont été obtenues, mais alors celles-ci n'avaient pas de sens puisque dans ces cas les variances augmentent rapidement, avec une tendance qui semble être exponentielle. Ceci se produit même pour des paramètres près de valeurs pour lesquelles les variances décroissent rapidement vers zéro.

On présente un exemple de ce phénomène à la figure 4.10, qui présente trois cas où l'on compare la solution du modèle de champ moyen de deuxième ordre (4.55) à la solution du modèle de champ moyen (4.17) pour les mêmes paramètres et la même condition initiale, à l'exception de l'activité entrante externe Q que l'on fait varier entre les trois cas. On a choisi ici de prendre pour les seuils une loi logistique, avec une fonction de répartition de la forme (4.31a). Les paramètres choisis ici sont

$$\alpha = 0,24 \text{ ms}^{-1}, \quad \beta = 1 \text{ ms}^{-1}, \quad \gamma = 1,8 \text{ ms}^{-1}, \quad \theta = 0, \quad s_\theta = 0,1, \quad c = 8, \quad (4.121)$$

avec l'état initial

$$(\mathcal{A}_0, \mathcal{R}_0, C_{AA0}, C_{RR0}, C_{AR0}) := (0,12; 0,49; 0,016; 0,095; -0,03), \quad (4.122)$$

où l'on ignore simplement les covariances pour obtenir la condition initiale à considérer pour le modèle de champ moyen.

La figure 4.10a illustre le cas où $Q = 0$. On voit bien que les variances décroissent rapidement vers zéro et que les solutions stables prévues par les deux modèles sont les mêmes. Toutefois, en ajoutant une petite inhibition externe $Q = -1/2$ (figure 4.10b), les variances se mettent à croître très rapidement pour sortir du domaine permis. Finalement, en augmentant l'inhibition externe jusqu'à $Q = -3/2$ (figure 4.10c), les solutions redeviennent très similaires.

Il est intéressant de comparer ces solutions à des trajectoires possibles du processus stochastique sous-jacent, illustrées à la figure 4.11. En comparant avec la figure 4.10, on constate que, du moins pour le cas $Q = -1/2$, la prédiction du modèle issu de l'approximation de champ moyen est beaucoup plus près des trajectoires du processus stochastique que la prédiction du modèle (4.55), et ce même si l'on ignore les variances.

On peut aussi tenter une approche plus systématique en vue de trouver des états stables avec des covariances non nulles. Pour ce faire, on se base sur les points fixes de formes *ii*) et *iii*) discutés au lemme 4.5. Si une solution débute près d'un tel point fixe, elle convergera vers lui s'il est stable, ce qui mènera à un état stable où les covariances sont non nulles. Sinon, à la lumière des résultats précédents, on s'attend à ce qu'elle converge plutôt vers un état stable avec des covariances nulles ou à ce qu'elle sorte du domaine physiologique, bien que d'autres comportements soient également possibles.

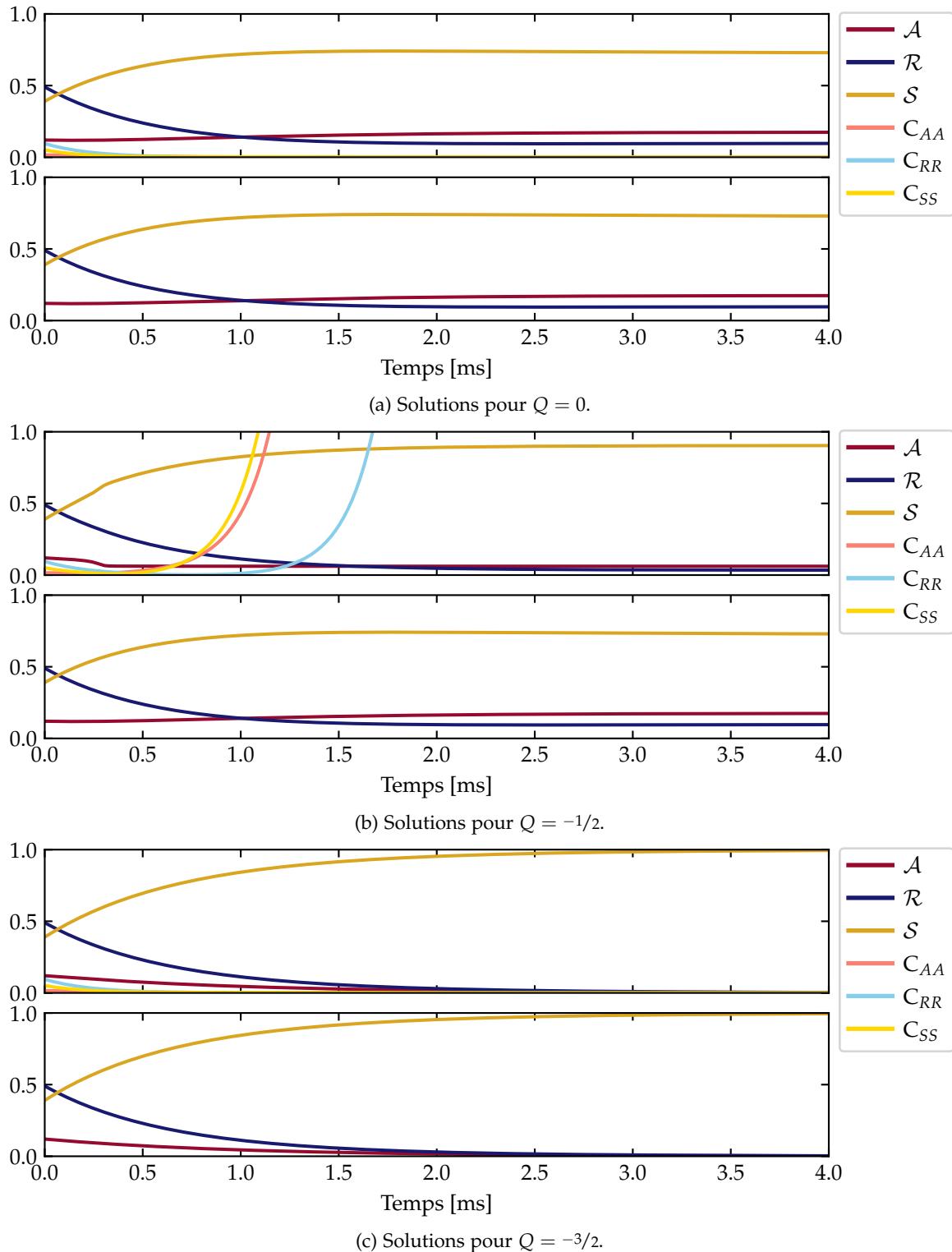


FIGURE 4.10 – Pour trois valeurs d’activité entrante Q , solutions des systèmes (4.55) (en haut) et (4.17) (en bas) pour les paramètres (4.121) et l’état initial (4.122).

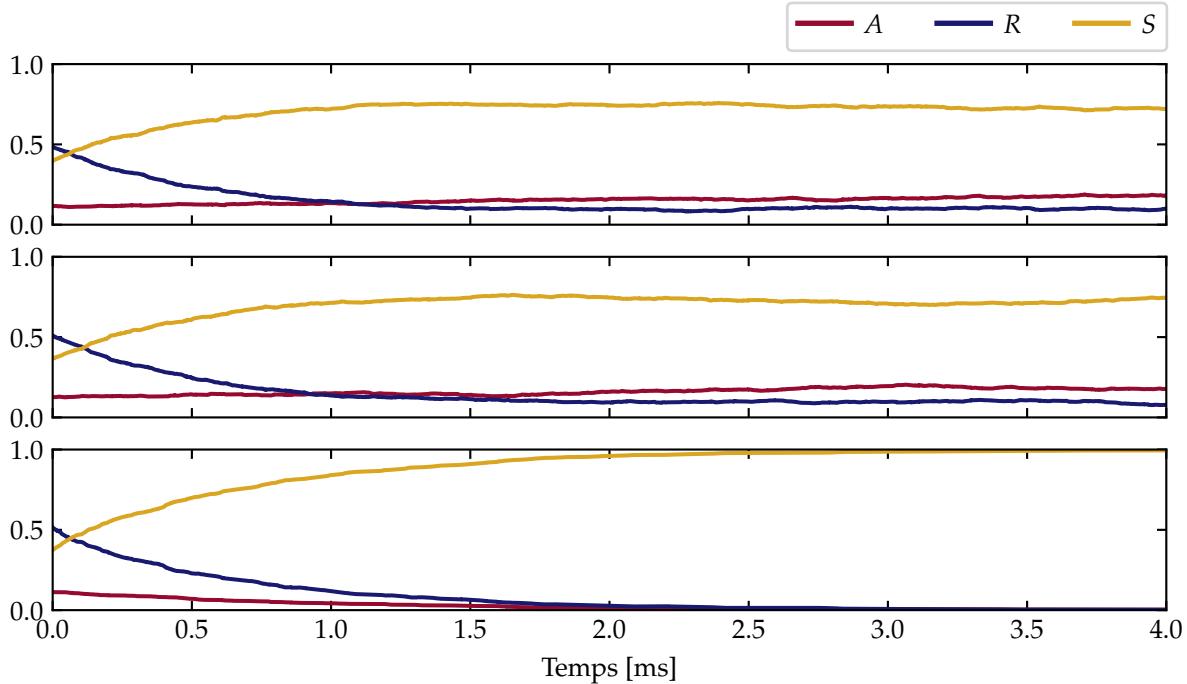


FIGURE 4.11 – Trajectoires possibles du processus de Markov construit à la section 4.2 pour les paramètres (4.121) et l'état initial (4.122), pour un réseau de 1000 neurones et avec $Q = 0$ en haut, $Q = -1/2$ au milieu et $Q = -3/2$ en bas. Voir l'annexe B.2 pour les détails sur la méthodologie employée pour les simulations.

Ainsi, on a voulu tester l'existence et la stabilité des points fixes de formes *ii)* et *iii)* en tentant d'explorer les domaines des paramètres. Pour ce faire, pour chacune des deux formes de points fixes, on a effectué un million de fois le test décrit par les étapes suivantes :

- i) Prendre des paramètres aléatoirement. Spécifiquement, on a pris les taux α , β et γ aléatoirement selon des lois de probabilité exponentielles de moyenne 5, et le coefficient de connexion c selon une loi de probabilité uniforme entre -10 et 10 . On a fixé le seuil moyen à 0 et son facteur d'échelle à $1/10$. Pour mesurer le temps en unités de $1/\beta$, on a ensuite divisé les taux α , β et γ par β .
- ii) Choisir la condition initiale de façon à la fixer au point fixe selon (4.117), et fixer la valeur de Q selon (4.116).
- iii) Si la condition initiale se trouve en dehors de \mathcal{D}_1^+ , rejeter ce cas-ci et passer au suivant.
- iv) Ajouter à la condition initiale une perturbation choisie aléatoirement sur une sphère de rayon 10^{-6} .
- v) Intégrer numériquement le système dynamique (4.55) sur 1000 unités de temps.
- vi) Si la solution est sortie de \mathcal{D}_1^+ , rejeter ce cas-ci et passer au suivant.
- vii) Si la norme du vecteur composé des trois covariances est inférieure à un certain seuil, choisi comme 10^{-5} , rejeter ce cas-ci et passer au suivant.
- viii) Reprendre les étapes *v)* à *vii)* en intégrant plutôt sur 100000 unités de temps.

Évidemment, plus on effectue de tests, plus on explore les valeurs possibles des paramètres,

mais plus il faut de temps pour réaliser les tests. Le choix — certes plutôt arbitraire — d'effectuer un million de tests pour chaque forme de points fixes est fait dans l'idée de trouver un équilibre entre la quantité de tests et leur durée. Le tout prend environ six heures avec un ordinateur portable standard, doté d'un processeur de 2,40 GHz et de 8 Gb de mémoire vive.

Au bout du compte, tous les cas testés finissent par être rejetés. Ces résultats suggèrent alors que les points fixes de forme *ii)* et *iii)* sont toujours instables, au moins lorsqu'ils se trouvent dans le domaine physiologique et que leurs composantes associées aux covariances sont assez grandes.

Les résultats numériques discutés dans cette sous-section suggèrent que, du moins dans le cas d'une seule population, le modèle de champ moyen de deuxième ordre (4.55) ne permet pas de représenter de nouveaux aspects de la dynamique réelle du graphe. Cependant, il reste pertinent de chercher des comportements intéressants avec davantage de populations.

§ 4.4.3. Exemple à deux populations

Lorsque le graphe est divisé en deux populations, le système dynamique (4.54) a 14 dimensions. Il devient donc difficile de l'étudier analytiquement. On présente donc ici un exemple étudié numériquement, trouvé alors qu'on cherchait à savoir si le système (4.54) permettait de prévoir des états stables distincts de ceux prévus par les modèles de champ moyen.

Les paramètres sont choisis de façon similaire à ceux choisis dans l'exemple étudié à la section 4.3.3. On considère deux populations notées E et I qui réfèrent respectivement à des populations excitatrice et inhibitrice. On suppose que les seuils suivent des lois logistiques dans ces deux populations, de sorte que les fonctions de répartition F_{θ_j} soient données par (4.31a). On fixe ensuite les paramètres

$$\alpha_E = 10 \text{ ms}^{-1}, \quad \beta_E = 4/5 \text{ ms}^{-1}, \quad \gamma_E = 4 \text{ ms}^{-1}, \quad \theta_E = 0, \quad s_{\theta_E} = 2/5, \quad Q_E = 0, \quad (4.123a)$$

$$\alpha_I = 9 \text{ ms}^{-1}, \quad \beta_I = 1 \text{ ms}^{-1}, \quad \gamma_I = 1 \text{ ms}^{-1}, \quad \theta_I = 6, \quad s_{\theta_I} = 2/5, \quad Q_I = 0 \quad (4.123b)$$

et la matrice de connexion

$$c = \begin{pmatrix} 12 & -16 \\ 14 & -4 \end{pmatrix}, \quad (4.123c)$$

et on choisit l'état initial

$$\begin{aligned} & (\mathcal{A}_E; \mathcal{A}_I; \mathcal{R}_E; \mathcal{R}_I; C_{AA}^{EE}; C_{AA}^{EI}; C_{AA}^{II}; C_{RR}^{EE}; C_{RR}^{EI}; C_{RR}^{II}; C_{AR}^{EE}; C_{AR}^{EI}; C_{AR}^{IE}; C_{AR}^{II}) \\ & = (0,4; 0,4; 0,08; 0,4; 0,073; 0,037; 0,019; 0,003; 0,008; 0,02; 0,016; 0,038; 0,008; 0,02). \end{aligned} \quad (4.124)$$

À partir de ces paramètres, on trouve la solution de l'équation différentielle (4.54) par une intégration numérique en commençant par la condition initiale (4.124). Il est alors intéressant de comparer le résultat au résultat de l'intégration numérique de l'équation différentielle (4.17) issue de l'approximation de champ moyen, en gardant les mêmes paramètres.

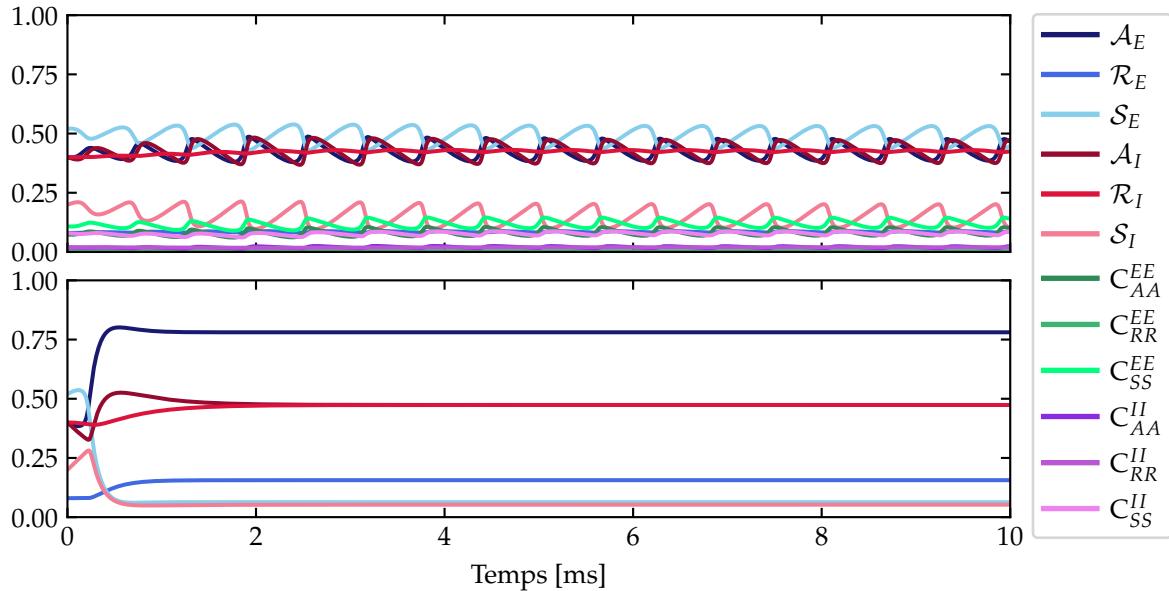


FIGURE 4.12 – En haut, solution de (4.54) pour les paramètres (4.123) et la condition initiale (4.124). En bas, solution de (4.17) pour les mêmes paramètres et la condition initiale (4.124) sans les covariances.

On doit alors adapter la condition initiale (4.124), puisque les covariances ne sont pas incluses dans le système (4.17). Pour ce faire, on néglige simplement les covariances, ce qui est équivalent à trouver la solution de (4.54) en prenant de (4.124) les composantes \mathcal{A}_J et \mathcal{R}_J de l'état initial et en fixant toutes les covariances à zéro.

Les résultats de ces deux intégrations numériques sont présentés à la figure 4.12. Les deux modèles prévoient ainsi des comportements très différents pour les deux cas : alors que le modèle de champ moyen prévoit que la solution converge vers un point fixe stable, le modèle qui inclut les covariances prévoit une convergence vers un cycle limite stable.

Ces résultats sont à première vue encourageants : ils suggèrent que l'extension (4.54) du modèle de champ moyen (4.17) permet, en incluant les covariances, de prédire des phénomènes différents qui seraient plus près de la réalité. Un tel résultat serait nouveau et montrerait que le système (4.54) est une meilleure approximation du processus stochastique sous-jacent que le système (4.17).

Pour s'en assurer, on compare les solutions obtenues à des trajectoires du processus stochastique, qui peuvent être obtenues par des simulations numériques. Un exemple d'une trajectoire typique est présentée à la figure 4.13. Malheureusement, il est alors clair que l'oscillation prévue par le modèle (4.54) ne représente pas du tout le comportement du processus stochastique, et que celui-ci correspond en fait beaucoup mieux à la solution du modèle de champ moyen.

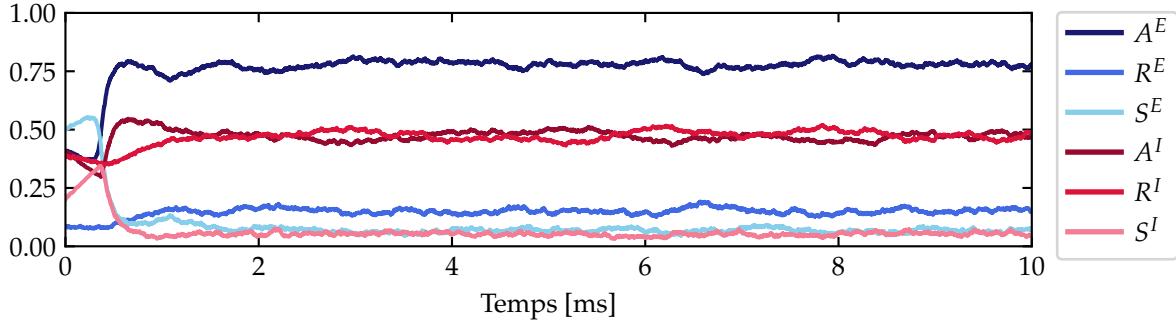


FIGURE 4.13 – Trajectoire possible du processus de Markov construit à la section 4.2 pour les paramètres (4.123) et l'état initial (4.124), pour un réseau de 1000 neurones par population. Voir l'annexe B.2 pour les détails sur la méthodologie employée pour les simulations.

§ 4.4.4. Conclusion

Si le modèle issu de l'approximation de champ moyen de deuxième ordre proposé à la section 4.4.1 semblait prometteur, on n'a finalement trouvé aucune façon de montrer qu'il permettait de modéliser davantage d'aspects de la dynamique initiale du graphe que le modèle issu de l'approximation de champ moyen.

En effet, même si on a pu trouver des points fixes dont les composantes associées aux covariances pouvaient être non nulles, on n'a pas pu montrer que ces points fixes pouvaient être stables. Par ailleurs, les résultats numériques discutés dans cette section ont permis de voir seulement des cas où les résultats du modèles qui inclut les covariances sont les mêmes que ceux du modèle de champ moyen, des cas où la prédiction du modèle qui inclut les covariances n'a pas de sens, ou bien des cas où les prédictions des deux modèles sont distinctes, mais où les résultats de simulations du processus stochastique sont plus près de la prédiction du modèle de champ moyen que de celle du modèle qui inclut les covariances.

Il semble donc approprié de conclure que l'approximation de champ moyen de deuxième ordre ne mène pas à un meilleur modèle de la dynamique macroscopique du réseau de neurones que l'approximation de champ moyen.

§ 4.5. Nouvelle solution au problème de fermeture de moments

On s'intéresse ici au système (4.16)–(4.42), donné par

$$\dot{\mathcal{A}}_J(t) = -\beta_J \mathcal{A}_J(t) + \alpha_J \mathbb{E}[S_t^J F_{\theta_J}(B_t^J)], \quad (4.16a)$$

$$\dot{\mathcal{R}}_J(t) = -\gamma_J \mathcal{R}_J(t) + \beta_J \mathcal{A}_J(t), \quad (4.16b)$$

$$\dot{C}_{AA}^{JK}(t) = -(\beta_J + \beta_K) C_{AA}^{JK}(t) + \alpha_K \text{Cov}[A_t^J, S_t^K F_{\theta_K}(B_t^K)] + \alpha_J \text{Cov}[S_t^J F_{\theta_J}(B_t^J), A_t^K], \quad (4.42a)$$

$$\dot{C}_{RR}^{JK}(t) = -(\gamma_J + \gamma_K) C_{RR}^{JK}(t) + \beta_K C_{AR}^{KJ}(t) + \beta_J C_{AR}^{JK}(t), \quad (4.42b)$$

$$\dot{C}_{AR}^{JK}(t) = -(\beta_J + \gamma_K) C_{AR}^{JK}(t) + \beta_K C_{AA}^{JK}(t) + \alpha_J \text{Cov}[S_t^J F_{\theta_J}(B_t^J), R_t^K]. \quad (4.42c)$$

On poursuit la recherche d'une solution au problème de fermeture de moments qui irait au-delà de l'approximation de champ moyen. À la section précédente, on a introduit une approximation de champ moyen de deuxième ordre, mais on a vu qu'elle ne permettait pas de mieux modéliser le comportement du processus stochastique sous-jacent. En effet, on a vu que le système (4.54) pouvait mener à des solutions qui sortent du domaine valide.

Intuitivement, on peut facilement penser que la faiblesse principale de l'approximation de champ moyen de deuxième ordre est qu'elle approxime des fonctions bornées par des fonctions qui se sont pas bornées. Par exemple, on a vu en (4.46) l'approximation

$$\mathbb{E}[S_t^J F_{\theta_J}(B_t^J)] \approx S_J(t) F_{\theta_J}(\mathcal{B}_J(t)) + F'_{\theta_J}(\mathcal{B}_J(t)) C_{SB}^{JJ}(t) + \frac{1}{2} S_J(t) F''_{\theta_J}(\mathcal{B}_J(t)) C_{BB}^{JJ}(t). \quad (4.46)$$

Or, la variable aléatoire S_t^J est bornée entre 0 et 1, de même que F_{θ_J} , puisqu'elle est une fonction de répartition. Ainsi, le membre de gauche de (4.46) est borné entre 0 et 1. Cependant, le membre de droite ne l'est pas en général. Le même phénomène se produit pour les autres approximations qui mènent au système dynamique (4.54) proposé à la section précédente. Ceci pourrait expliquer les cas où les solutions du système (4.54) sortent du domaine physiologique, comme celui présenté à la figure 4.10.

On propose alors dans cette section une autre solution au problème de fermeture de moments, qui est construite à partir de fonctions bornées. Puisque l'essentiel du défi provient des fonctions de répartition des seuils, on se concentre d'abord sur la construction d'une fonction qu'on utilisera afin d'approximer l'espérance de fonctions de forme sigmoïdale. En supposant que les seuils suivent des lois de probabilité unimodales et symétriques, on utilise cette fonction pour proposer une nouvelle solution au problème de fermeture de moments pour le système (4.16)–(4.42). On présente finalement quelques exemples où ce nouveau système est comparé au système issu de l'approximation de champ moyen.

§ 4.5.1. Approximation de l'espérance d'une fonction de forme sigmoïdale

On veut d'abord approximer l'espérance de la fonction de répartition des seuils F_{θ_J} dans une population J lorsqu'elle est évaluée à l'activité entrante B_t^J . On suppose ici que la loi de probabilité des seuils dans la population J est unimodale et symétrique et que F_{θ_J} est quelques fois dérivable, de sorte que F_{θ_J} ait une forme sigmoïdale.

Afin de considérer les covariances, il est logique de rechercher une approximation de la forme $\mathbb{E}[F_{\theta_J}(B_t^J)] \approx G_J(\mathcal{B}_J(t), C_{BB}^{JJ}(t))$, où $G_J: \mathbb{R} \times [0, \infty) \rightarrow [0, 1]$ est à déterminer.

§ 4.5.1.1. Conditions à respecter

On obtient d'abord quelques caractéristiques que doit posséder G_J pour correspondre à l'intuition sur ce que représente $\mathbb{E}[F_{\theta_J}(B_t^J)]$. D'abord, si J reçoit une plus grande activité entrante, elle répondra davantage. Ainsi, G_J doit être croissante selon son premier argu-

ment. Ensuite, si J reçoit une très grande inhibition elle ne devrait pas répondre, alors que si elle reçoit une très grande excitation, sa réponse devrait être maximale. Autrement dit, $G_J(\mathcal{B}_J, C_{BB}^{JJ})$ devrait tendre vers 0 lorsque $\mathcal{B}_J \rightarrow -\infty$, et vers 1 lorsque $\mathcal{B}_J \rightarrow \infty$.

On a également un peu d'information sur la variable aléatoire $F_{\theta_J}(B_t^J)$ pour des valeurs extrêmes de la variance de B_t^J . Si cette variance est nulle, alors $F_{\theta_J}(B_t^J) = F_{\theta_J}(\mathcal{B}_J(t))$ p.s., donc on veut que $G_J(\mathcal{B}_J, 0) = F_{\theta_J}(\mathcal{B}_J)$ pour n'importe quelle valeur moyenne \mathcal{B}_J . Au contraire, si la variance de B_t^J est très grande, $F_{\theta_J}(B_t^J)$ devrait tendre vers une variable aléatoire binaire qui vaut 0 ou 1 avec probabilités $1/2$, et ce peu importe la valeur de \mathcal{B}_J . Ainsi, on veut que $G_J(\mathcal{B}_J, C_{BB}^{JJ}) \rightarrow 1/2$ lorsque $C_{BB}^{JJ} \rightarrow \infty$.

On sait ensuite que l'approximation de Taylor d'ordre 2 de F_{θ_J} autour de $\mathcal{B}_J(t)$ devrait être valide pour de très petites variances de B_t^J . Cette approximation dicte que

$$\begin{aligned}\mathbb{E}[F_{\theta_J}(B_t^J)] &\approx \mathbb{E}\left[F_{\theta_J}(\mathcal{B}_J(t)) + F'_{\theta_J}(\mathcal{B}_J(t))(B_t^J - \mathcal{B}_J(t)) + \frac{1}{2}F''_{\theta_J}(\mathcal{B}_J(t))(B_t^J - \mathcal{B}_J(t))^2\right] \\ &= F_{\theta_J}(\mathcal{B}_J(t)) + \frac{1}{2}F''_{\theta_J}(\mathcal{B}_J(t))C_{BB}^{JJ}(t).\end{aligned}\quad (4.125)$$

Celle-ci doit être bonne pour toute valeur de $\mathcal{B}_J(t)$ lorsque $C_{BB}^{JJ}(t)$ est petite, donc la dérivée de G_J par rapport à son second argument doit être $\frac{1}{2}F''_{\theta_J}(\mathcal{B}_J(t))$ lorsque évaluée à $(\mathcal{B}_J(t), 0)$.

Finalement, l'étude du système résultant sera plus difficile si G_J n'est pas suffisamment régulière. On cherche donc une approximation qui soit continûment dérivable.

En résumé, on cherche $G_J: \mathbb{R} \times [0, \infty) \rightarrow [0, 1]$ telle que :

- i) G_J est de classe \mathcal{C}^1 sur $\mathbb{R} \times [0, \infty)$;
- ii) $\forall x \in \mathbb{R}, G_J(x, 0) = F_{\theta_J}(x)$ et $G(x, y) \rightarrow 1/2$ lorsque $y \rightarrow \infty$;
- iii) $\forall y \geq 0, G_J(x, y) \rightarrow 0$ lorsque $x \rightarrow -\infty$ et $G_J(x, y) \rightarrow 1$ lorsque $x \rightarrow \infty$;
- iv) $\forall y \geq 0$, la fonction $x \mapsto G_J(x, y)$ est croissante;
- v) $\forall x \in \mathbb{R}, \partial_y G_J(x, 0) = \frac{1}{2}F''_{\theta_J}(x)$.

§ 4.5.1.2. Construction de l'approximation

Afin de trouver une fonction qui respecte les conditions données à la section précédente, il est instructif de débuter par trouver une intuition de ce à quoi pourrait ressembler l'espérance de $F_{\theta_J}(B_t^J)$ lorsque B_t^J a une variance non nulle. Dans ce contexte, il est plus simple de travailler avec une forme explicite pour F_{θ_J} , afin de pouvoir observer la situation graphiquement. On suppose donc ici que les seuils suivent une loi logistique dans la population J . On rappelle que dans ce cas, la fonction de répartition $F_{\theta_J}: \mathbb{R} \rightarrow (0, 1)$ est donnée par

$$F_{\theta_J}(x) := \frac{1}{1 + \exp\left(\frac{x - \theta_J}{s_{\theta_J}}\right)}, \quad (4.126)$$

où s_{θ_J} est un facteur d'échelle qui correspond à l'écart type des seuils à un facteur $\pi/\sqrt{3}$ près. Un exemple de cette fonction est donné à la figure 4.14.

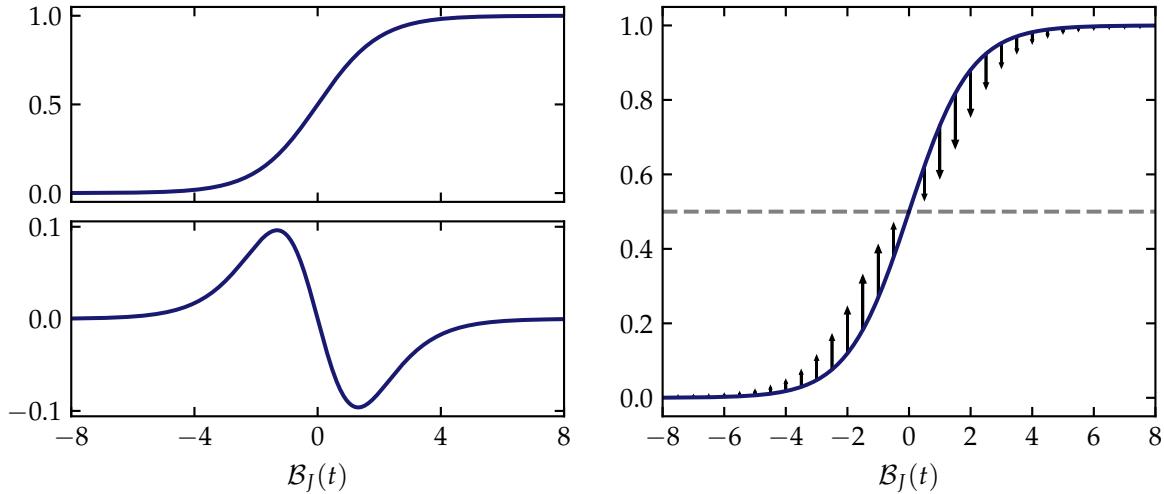


FIGURE 4.14 – En haut à gauche, exemple de graphe de F_{θ_J} si les seuils suivent une loi logistique de moyenne $\theta_J = 0$ et de facteur d'échelle $s_{\theta_J} = 1$. En bas à gauche, graphe de la dérivée seconde de la fonction F_{θ_J} donnée en haut. À droite, schématisation d'une estimation de l'espérance de $F_{\theta_J}(B_t^J)$ selon la valeur de $B_J(t)$. Chaque flèche représente la translation d'une valeur de $F_{\theta_J}(B_J(t))$ vers la valeur estimée de l'espérance de $F_{\theta_J}(B_t^J)$ lorsque l'espérance de B_t^J est $B_J(t)$.

Pour obtenir une intuition sur $\mathbb{E}[F_{\theta_J}(B_t^J)]$, on imagine que B_t^J suive une loi normale de moyenne $B_J(t)$ et d'écart type égal à 1. Dans ce cas, la probabilité que B_t^J se trouve dans un intervalle d'une largeur de 4 centré sur $B_J(t)$ est d'environ 95%. On peut alors facilement imaginer si $F_{\theta_J}(B_J(t))$ surestime ou sous-estime la valeur de $\mathbb{E}[F_{\theta_J}(B_t^J)]$ pour différentes valeurs de $B_J(t)$, en imaginant que toutes les valeurs possibles de B_t^J se trouvent dans l'intervalle $[B_J(t) - 2, B_J(t) + 2]$.

En observant la figure 4.14, on voit que la valeur de $F_{\theta_J}(B_t^J)$ ne varie pas beaucoup si B_t^J se trouve dans l'intervalle $[-8, -4]$, qui correspond à une valeur moyenne de $B_J(t) = -6$. Ainsi, à ce point, $F_{\theta_J}(B_J(t))$ devrait être une bonne approximation de $\mathbb{E}[F_{\theta_J}(B_t^J)]$. Déplaçant l'intervalle de 2 vers la droite, les valeurs de $F_{\theta_J}(B_t^J)$ sont moins uniformes, étant donnée l'augmentation de F_{θ_J} vers la droite de l'intervalle. Dans ce cas, la valeur de $F_{\theta_J}(B_J(t))$ risque d'être *plus faible* que $\mathbb{E}[F_{\theta_J}(B_t^J)]$. Cet effet s'accentue lorsqu'on déplace à nouveau l'intervalle de 2 vers la droite. Cependant, il devient moins important en approchant trop le centre de l'intervalle de zéro, qui correspond au seuil moyen, car F_{θ_J} devient alors de plus en plus antisymétrique dans l'intervalle. À l'extrême, pour un intervalle centré à zéro, les valeurs de $F_{\theta_J}(B_t^J)$ sont réparties symétriquement autour de $1/2$, de sorte que $\mathbb{E}[F_{\theta_J}(B_t^J)] = 1/2 = F_{\theta_J}(0) = F_{\theta_J}(\theta_J)$.

Comme F_{θ_J} est impaire selon l'axe $y = 1/2$ autour de θ_J , c'est-à-dire que $F_{\theta_J}(\theta_J - x) = 1 - F_{\theta_J}(\theta_J + x)$ pour chaque $x \in \mathbb{R}$, les remarques précédentes s'appliquent pour des valeurs moyennes $B_J(t)$ positives, mais à l'inverse. Ainsi, lorsque $B_J(t)$ dépasse zéro, on peut d'abord s'attendre à ce que $F_{\theta_J}(B_J(t))$ surestime la valeur de $\mathbb{E}[F_{\theta_J}(B_t^J)]$, et lorsque $B_J(t)$ de-

vient trop élevée et que $F_{\theta_J}(B_t^J)$ ne varie presque plus, cet effet est très atténué et $F_{\theta_J}(\mathcal{B}_J(t))$ risque d'être une bonne estimation de $\mathbb{E}[F_{\theta_J}(B_t^J)]$.

La discussion précédente est schématisée à droite de la figure 4.14. Il est intéressant de remarquer que, d'après la forme de F''_{θ_J} , si la variance de B_t^J est non nulle, mais petite, alors son effet sur $\mathbb{E}[F_{\theta_J}(B_t^J)]$ prévu par la discussion précédente correspond à celui prévu par l'approximation de Taylor (4.125). Globalement, la translation de la courbe de F_{θ_J} vers ce à quoi pourrait ressembler $\mathbb{E}[F_{\theta_J}(B_t^J)]$, qui est décrite par les flèches sur la figure 4.14, s'apparente à une augmentation de l'écart type de la loi des seuils.

Une fois l'écart type de la loi des seuils augmenté, la nouvelle fonction de répartition peut facilement être exprimée en fonction de l'ancienne. En effet, si l'on note θ_j la variable aléatoire qui représente le seuil d'un neurone $j \in J$ et si l'on note σ_{θ_j} son écart type, alors augmenter l'écart type de la loi comme $\sigma_{\theta_j} \mapsto (1+a)\sigma_{\theta_j}$ pour un $a \geq 0$ correspond à modifier la variable aléatoire θ_j comme $\theta_j \mapsto \theta_J + (\theta_j - \theta_J)(1+a) = \theta_j(1+a) - a\theta_J$. Or,

$$\mathbb{P}\left[\theta_j(1+a) - a\theta_J \geq x\right] = \mathbb{P}\left[\theta_j \geq \frac{x + a\theta_J}{1+a}\right] = F_{\theta_J}\left(\frac{x + a\theta_J}{1+a}\right), \quad (4.127)$$

où F_{θ_J} note toujours la fonction de répartition de véritable la loi de probabilité des seuils. Ainsi, un changement d'écart type $\sigma_{\theta_j} \mapsto (1+a)\sigma_{\theta_j}$ correspond à un changement de la fonction de répartition F_{θ_J} vers la fonction $x \mapsto F_{\theta_J}\left(\frac{x + a\theta_J}{1+a}\right)$.

On suppose que le changement d'écart type recherché peut dépendre à la fois de la valeur moyenne $\mathcal{B}_J(t)$ et de la variance de B_t^J . En effet, on s'attend à ce que l'effet soit négligeable pour une variance presque nulle, mais pas pour des variances plus élevées. On recherche donc une approximation G_J de la forme

$$G_J(x, y) := F_{\theta_J}\left(\frac{x + \theta_J g_J(x, y)}{1 + g_J(x, y)}\right), \quad (4.128)$$

où $g_J: \mathbb{R} \times [0, \infty) \rightarrow [0, \infty)$ est une fonction inconnue, suffisamment régulière.

Les conditions données à la section 4.5.1.1 permettent de trouver facilement un candidat pour g_J . En effet, on veut d'après la condition *ii*) que $G_J(x, 0) = F_{\theta_J}(x)$ pour tout x , ce qui est vrai si $g_J(x, 0) = 0$. On remarque ensuite que

$$\partial_y G_J(x, y) = F'_{\theta_J}\left(\frac{x + \theta_J g_J(x, y)}{1 + g_J(x, y)}\right) \left(\frac{\theta_J \partial_y g_J(x, y)}{1 + g_J(x, y)} - \frac{(x + \theta_J g_J(x, y)) \partial_y g_J(x, y)}{(1 + g_J(x, y))^2} \right) \quad (4.129)$$

$$= F'_{\theta_J}\left(\frac{x + \theta_J g_J(x, y)}{1 + g_J(x, y)}\right) \frac{\theta_J - x}{(1 + g_J(x, y))^2} \partial_y g_J(x, y). \quad (4.130)$$

Ainsi, avec $g_J(x, 0) = 0$,

$$\partial_y G_J(x, 0) = F'_{\theta_J}(x)(\theta_J - x) \partial_y g_J(x, 0). \quad (4.131)$$

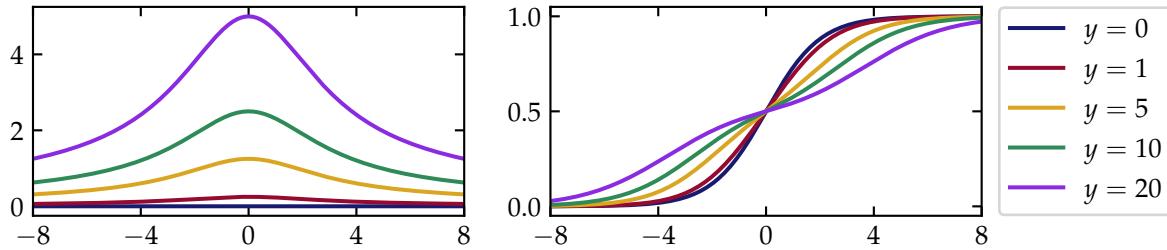


FIGURE 4.15 – À gauche, exemples de fonctions $x \mapsto g_J(x, y)$ avec g_J donnée par (4.133), en supposant que les seuils suivent une loi logistique dans J . À droite, fonctions $x \mapsto G_J(x, y)$ correspondantes d'après (4.128). On suppose pour ces exemples un seuil moyen $\theta_J = 0$ et un facteur d'échelle $s_{\theta_J} = 1$.

La condition *v)* impose que $\partial_y G_J(x, 0) = \frac{1}{2} F''_{\theta_J}(x)$. Pour qu'elle soit respectée, il est donc nécessaire que

$$\partial_y g_J(x, 0) = \frac{1}{2(\theta_J - x)} \frac{F''_{\theta_J}(x)}{F'_{\theta_J}(x)}. \quad (4.132)$$

Pour $x \neq \theta_J$, la dernière condition amène naturellement à prendre comme candidat

$$g_J(x, y) := \frac{y}{2(\theta_J - x)} \frac{F''_{\theta_J}(x)}{F'_{\theta_J}(x)}, \quad (4.133)$$

qui définit g_J sur $(\mathbb{R} \setminus \{\theta_J\}) \times [0, \infty)$ tant que $F'_{\theta_J} > 0$. Il restera à définir g_J en $x = \theta_J$.

Il est intéressant de remarquer que si les seuils suivent une loi logistique, cette fonction a même une forme de cloche selon son premier argument. Quelques exemples de fonctions $x \mapsto G_J(x, y)$ et $x \mapsto g_J(x, y)$ pour différentes valeurs de y sont données à la figure 4.15.

Sous certaines conditions supplémentaires, on peut montrer que la fonction proposée plus haut respecte bien les conditions données à la section 4.5.1.1. Pour le montrer, on commence par donner quelques résultats sur la fonction g_J .

Lemme 4.8. Soit $F_{\theta_J} : \mathbb{R} \rightarrow [0, 1]$ la fonction de répartition d'une loi de probabilité unimodale et symétrique autour de θ_J . On suppose que F_{θ_J} est de classe \mathcal{C}^3 et que $F'_{\theta_J} > 0$. Soit $g_J : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$g_J(x, y) := \begin{cases} \frac{y}{2(\theta_J - x)} \frac{F''_{\theta_J}(x)}{F'_{\theta_J}(x)} & \text{si } x \neq \theta_J, \\ -\frac{y}{2} \frac{F'''_{\theta_J}(\theta_J)}{F'_{\theta_J}(\theta_J)} & \text{si } x = \theta_J. \end{cases} \quad (4.134)$$

Alors :

- i) Pour tout $x \in \mathbb{R}$, $g_J(x, 0) = 0$;
- ii) $g_J \geq 0$ sur $\mathbb{R} \times [0, \infty)$;
- iii) g_J est continue.

◆

Démonstration. D'abord, le point *i)* est évident puisque g_J est toujours linéaire selon y . Ensuite, puisque la loi des seuils est unimodale et symétrique autour de θ_J , on sait d'après

la proposition 1.8 que $F''_{\theta_J}(x) \geqslant 0$ lorsque $x \leqslant \theta_J$. Ainsi, le rapport de $F''_{\theta_J}(x)$ et $\theta_J - x$ est toujours positif. Comme on suppose que $F_{\theta_J} > 0$, alors $g_J(x, y) \geq 0$ pour tout $x \neq \theta_J$ et pour tout $y \geq 0$. Finalement, puisque F''_{θ_J} est positive à gauche de θ_J mais négative à sa droite, elle doit être décroissante en θ_J . Ainsi, $F'''_{\theta_J}(\theta_J) \leq 0$ et $g_J(\theta_J, y) \geq 0$. Ceci termine de montrer *ii*).

Finalement, il est clair que g_J est continue sur $(\mathbb{R} \setminus \{\theta_J\}) \times \mathbb{R}$ et sur $\{\theta_J\} \times \mathbb{R}$. Il reste à montrer que pour tout $y \in \mathbb{R}$, g_J est bien continue en $x = \theta_J$. Comme $F''_{\theta_J}(\theta_J) = 0$, la règle de l'Hôpital implique que

$$\lim_{x \rightarrow \theta_J} \frac{F''_{\theta_J}(x)}{\theta_J - x} = \lim_{x \rightarrow \theta_J} \frac{F'''_{\theta_J}(x)}{-1} = -F'''_{\theta_J}(\theta_J). \quad (4.135)$$

Fixant alors $y \in \mathbb{R}$, le résultat précédent peut être utilisé pour calculer que

$$\lim_{x \rightarrow \theta_J} g_J(x, y) = \lim_{x \rightarrow \theta_J} \frac{y}{2(\theta_J - x)} \frac{F''_{\theta_J}(x)}{F'_{\theta_J}(x)} = \frac{y}{2F'_{\theta_J}(\theta_J)} \cdot -F'''_{\theta_J}(\theta_J) = g_J(\theta_J, y). \quad (4.136)$$

g_J est donc bien continue en $x = \theta_J$ pour tout y , ce qui termine de montrer *iii*). ■

Lemme 4.9. Soient F_{θ_J} et g_J les fonctions introduites au lemme 4.8. Si F_{θ_J} est de classe \mathcal{C}^4 , alors g_J est de classe \mathcal{C}^1 sur \mathbb{R}^2 . De plus,

$$\partial_x g_J(x, y) = \begin{cases} \frac{y}{2(\theta_J - x)F'_{\theta_J}(x)} \left(\frac{F''_{\theta_J}(x)}{\theta_J - x} + F'''_{\theta_J}(x) - \frac{(F''_{\theta_J}(x))^2}{F'_{\theta_J}(x)} \right) & \text{si } x \neq \theta_J, \\ 0 & \text{si } x = \theta_J, \end{cases} \quad (4.137)$$

et

$$\partial_y g_J(x, y) = \begin{cases} \frac{1}{2(\theta_J - x)F'_{\theta_J}(x)} \frac{F''_{\theta_J}(x)}{F'_{\theta_J}(x)} & \text{si } x \neq \theta_J, \\ -\frac{1}{2} \frac{F'''_{\theta_J}(\theta_J)}{F'_{\theta_J}(\theta_J)} & \text{si } x = \theta_J. \end{cases} \quad (4.138)$$

◆

Démonstration. On note d'abord que pour tout $x \in \mathbb{R}$, la fonction $y \mapsto g_J(x, y)$ est linéaire. Sa dérivée existe donc toujours, et un calcul direct mène à la formule (4.138). De plus, comparant avec la définition de g_J , on voit que $\partial_y g_J(x, y) = g_J(x, 1)$ pour tout $y \in \mathbb{R}$. Ainsi, puisque le lemme 4.8 montre que g_J est continue, alors $\partial_y g_J$ est continue.

Il reste à montrer l'énoncé correspondant pour la dérivée partielle $\partial_x g_J$. On remarque d'abord que d'après le lemme 4.8, la fonction $x \mapsto g(x, 0)$ est identiquement nulle sur \mathbb{R} . L'existence et la continuité de la dérivée de cette fonction sont donc triviales.

On fixe alors $y \neq 0$. Pour $x \neq \theta_J$, on calcule d'abord que

$$\partial_x g_J(x, y) = \frac{y}{2} \left(\frac{F'''_{\theta_J}(x)}{(\theta_J - x)F'_{\theta_J}(x)} + \frac{F''_{\theta_J}(x)}{(\theta_J - x)^2 F'_{\theta_J}(x)} - \frac{1}{\theta_J - x} \left(\frac{F''_{\theta_J}(x)}{F'_{\theta_J}(x)} \right)^2 \right), \quad (4.139)$$

d'où

$$\partial_x g_J(x, y) = \frac{y}{2(\theta_J - x)} \left(\frac{F_{\theta_J}'''(x)}{F_{\theta_J}'(x)} + \frac{F_{\theta_J}''(x)}{(\theta_J - x)F_{\theta_J}'(x)} - \left(\frac{F_{\theta_J}''(x)}{F_{\theta_J}'(x)} \right)^2 \right). \quad (4.140)$$

Puisque $F_{\theta_J}' > 0$, cette expression est bien définie et continue pour tout $x \neq \theta_J$. Il reste à vérifier l'existence et la continuité de $\partial_x g_J$ en $x = \theta_J$. Pour $x \neq \theta_J$,

$$\frac{g_J(x, y) - g_J(\theta_J, y)}{x - \theta_J} = \frac{y}{2(x - \theta_J)} \left(\frac{F_{\theta_J}''(x)}{(\theta_J - x)F_{\theta_J}'(x)} + \frac{F_{\theta_J}'''(\theta_J)}{F_{\theta_J}'(\theta_J)} \right) \quad (4.141)$$

$$= -\frac{y}{2} \frac{F_{\theta_J}'(\theta_J)F_{\theta_J}''(x) + (\theta_J - x)F_{\theta_J}'''(\theta_J)F_{\theta_J}'(x)}{(\theta_J - x)^2 F_{\theta_J}'(\theta_J)F_{\theta_J}'(x)}. \quad (4.142)$$

Lorsque $x \rightarrow \theta_J$, le numérateur et le dénominateur de cette expression tendent tous deux vers 0. La règle de l'Hôpital montre donc que

$$\partial_x g_J(\theta_J, y) = \lim_{x \rightarrow \theta_J} \frac{g_J(x, y) - g_J(\theta_J, y)}{x - \theta_J} \quad (4.143)$$

$$= -\frac{y}{2} \lim_{x \rightarrow \theta_J} \frac{F_{\theta_J}'(\theta_J)F_{\theta_J}'''(x) - F_{\theta_J}'''(\theta_J)F_{\theta_J}'(x) + (\theta_J - x)F_{\theta_J}'''(\theta_J)F_{\theta_J}''(x)}{(\theta_J - x)^2 F_{\theta_J}'(\theta_J)F_{\theta_J}''(x) - 2(\theta_J - x)F_{\theta_J}'(\theta_J)F_{\theta_J}'(x)}. \quad (4.144)$$

Un calcul direct montre que le dernier terme tend vers zéro lorsque $x \rightarrow \theta_J$. En utilisant à nouveau la règle de l'Hôpital pour traiter les deux autres, on peut vérifier que

$$\partial_x g_J(\theta_J, y) = -\frac{y}{4} \frac{F_{\theta_J}'''(\theta_J)}{F_{\theta_J}'(\theta_J)}. \quad (4.145)$$

Par la proposition 1.8, on sait que F_{θ_J}''' est impaire autour de θ_J , de sorte que $F_{\theta_J}'''(\theta_J) = 0$. Ainsi, on voit que $\partial_x g_J(\theta_J, y) = 0$.

Il reste à vérifier la continuité de $\partial_x g_J$ pour $x = \theta_J$. Pour $x \neq \theta_J$, on a d'après (4.140) que

$$\partial_x g_J(x, y) = \frac{y}{2(\theta_J - x)} \left[\frac{F_{\theta_J}'''(x)}{F_{\theta_J}'(x)} - \frac{F_{\theta_J}'''(\theta_J)}{F_{\theta_J}'(\theta_J)} - \left(\frac{F_{\theta_J}''(x)}{F_{\theta_J}'(x)} \right)^2 + \frac{F_{\theta_J}''(x)}{(\theta_J - x)F_{\theta_J}'(x)} + \frac{F_{\theta_J}'''(\theta_J)}{F_{\theta_J}'(\theta_J)} \right]. \quad (4.146)$$

Comparant avec (4.141), on voit donc que

$$\begin{aligned} \partial_x g_J(x, y) &= \frac{y}{2} \frac{F_{\theta_J}'(\theta_J)F_{\theta_J}'''(x) - F_{\theta_J}'''(\theta_J)F_{\theta_J}'(x)}{(\theta_J - x)F_{\theta_J}'(\theta_J)F_{\theta_J}'(x)} \\ &\quad - \frac{y}{2} \frac{F_{\theta_J}''(x)}{\theta_J - x} \frac{F_{\theta_J}''(x)}{(F_{\theta_J}'(x))^2} - \frac{g_J(x, y) - g_J(\theta_J, y)}{x - \theta_J}. \end{aligned} \quad (4.147)$$

Avec les propriétés déjà utilisées jusqu'ici, on peut alors facilement vérifier que les deux premiers termes tendent vers zéro lorsque $x \rightarrow \theta_J$. On obtient donc que

$$\lim_{x \rightarrow \theta_J} \partial_x g_J(x, y) = -\partial_x g_J(\theta_J, y) = 0. \quad (4.148)$$

Ainsi, la dérivée partielle $\partial_x g_J$ est continue partout sur \mathbb{R}^2 .

Comme les deux dérivées partielles $\partial_x g_J$ et $\partial_y g_J$ existent et sont continues partout sur \mathbb{R}^2 , on voit que g_J est bien de classe \mathcal{C}^1 sur ce domaine. ■

On peut maintenant montrer que, sous quelques suppositions supplémentaires sur F_{θ_J} et g_J , $G_J(B_J(t), C_{BB}^{JJ}(t))$ est effectivement une bonne approximation de $\mathbb{E}[F_{\theta_J}(B_t^J)]$ au sens où elle respecte les conditions données à la section 4.5.1.1.

Théorème 4.10. Soit $G_J : \mathbb{R}^2 \rightarrow [0, 1]$ définie par

$$G_J(x, y) := F_{\theta_J}\left(\frac{x + \theta_J g_J(x, y)}{1 + g_J(x, y)}\right), \quad (4.149)$$

où g_J est définie de même qu'au lemme 4.8. En plus des suppositions tenues sur F_{θ_J} au lemme 4.8, on suppose que :

- i) F_{θ_J} est de classe \mathcal{C}^4 sur \mathbb{R} ;
- ii) $\forall y \geq 0$, la fonction $x \mapsto g_J(x, y)$ est bornée sur \mathbb{R} ;
- iii) $\forall x \in \mathbb{R}$ et $\forall y \geq 0$, $1 + g_J(x, y) + (\theta_J - x) \partial_x g_J(x, y) \geq 0$.

Alors, G_J respecte les conditions suivantes :

- i) G_J est de classe \mathcal{C}^1 sur \mathbb{R}^2 ;
- ii) $\forall x \in \mathbb{R}$, $G_J(x, 0) = F_{\theta_J}(x)$ et $G(x, y) \rightarrow 1/2$ lorsque $y \rightarrow \infty$;
- iii) $\forall y \geq 0$, $G_J(x, y) \rightarrow 0$ lorsque $x \rightarrow -\infty$ et $G_J(x, y) \rightarrow 1$ lorsque $x \rightarrow \infty$;
- iv) $\forall y \geq 0$, la fonction $x \mapsto G_J(x, y)$ est croissante;
- v) $\forall x \in \mathbb{R}$, $\partial_y G_J(x, 0) = \frac{1}{2} F_{\theta_J}''(x)$. ◆

Démonstration. Par le lemme 4.9, g_J est de classe \mathcal{C}^1 sur \mathbb{R}^2 . La fonction $(x, y) \mapsto \frac{x + \theta_J y}{1 + y}$ l'est aussi. Comme G_J est issue de leur composition, elle l'est donc aussi. Ceci montre i).

On étudie ensuite les comportements limites de G_J selon y . On fixe donc $x \in \mathbb{R}$. D'après le lemme 4.8, on sait que $g_J(x, 0) = 0$. Ainsi, $G_J(x, 0) = F_{\theta_J}(x)$. De plus, puisque $g_J(x, 1)$ est positive, on voit directement que $g_J(x, y) = yg_J(x, 1) \rightarrow \infty$ lorsque $y \rightarrow \infty$, ce qui implique que $(x + \theta_J g_J(x, y)) / (1 + g_J(x, y)) \rightarrow \theta_J$ lorsque $y \rightarrow \infty$. Ainsi, on trouve que $G_J(x, y) \rightarrow F_{\theta_J}(\theta_J) = 1/2$ lorsque $y \rightarrow \infty$. Ceci termine de montrer ii).

On passe ensuite aux comportements limites de G_J selon x . On fixe donc $y \geq 0$. On sait que la fonction $x \mapsto g_J(x, y)$ est bornée sur \mathbb{R} , de sorte que

$$\lim_{x \rightarrow \pm\infty} \frac{x + \theta_J g_J(x, y)}{1 + g_J(x, y)} = \pm\infty. \quad (4.150)$$

On obtient donc que

$$\lim_{x \rightarrow -\infty} G_J(x, y) = \lim_{x \rightarrow -\infty} F_{\theta_J}(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow \infty} G_J(x, y) = \lim_{x \rightarrow \infty} F_{\theta_J}(x) = 1, \quad (4.151)$$

ce qui montre iii).

On sait ensuite que la fonction $x \mapsto G_J(x, 0)$ est croissante, puisqu'elle toujours égale à F_{θ_J} , qui est une fonction de répartition. Fixant $y > 0$, on calcule aisément que

$$\partial_x G_J(x, y) = F'_{\theta_J}\left(\frac{x + \theta_J g_J(x, y)}{1 + g_J(x, y)}\right) \frac{1 + g_J(x, y) + (\theta_J - x) \partial_x g_J(x, y)}{(1 + g_J(x, y))^2}. \quad (4.152)$$

On sait que $F'_{\theta_J} > 0$, et on suppose que $1 + g_J(x, y) + (\theta_J - x) \partial_x g_J(x, y) \geq 0$. Ainsi, on voit que $\partial_x G_J(x, y) \geq 0$ pour tout $x \in \mathbb{R}$, c'est-à-dire que la fonction $x \mapsto G_J(x, y)$ est croissante. Ceci termine de montrer *iv*).

Finalement, un calcul direct permet de voir que

$$\partial_y G_J(x, y) = F'_{\theta_J} \left(\frac{x + \theta_J g_J(x, y)}{1 + g_J(x, y)} \right) \frac{(\theta_J - x) \partial_y g_J(x, y)}{(1 + g_J(x, y))^2}. \quad (4.153)$$

On sait que $g_J(x, 0) = 0$ pour tout $x \in \mathbb{R}$. Ainsi,

$$\partial_y G_J(x, 0) = F'_{\theta_J}(x)(\theta_J - x) \partial_y g_J(x, 0). \quad (4.154)$$

À partir de l'expression de $\partial_y g_J$ donnée au lemme 4.9, on trouve donc que pour $x \neq \theta_J$,

$$\partial_y G_J(x, 0) = F'_{\theta_J}(x)(\theta_J - x) \frac{1}{2(\theta_J - x)} \frac{F''_{\theta_J}(x)}{F'_{\theta_J}(x)} = \frac{1}{2} F''_{\theta_J}(x). \quad (4.155)$$

De plus, il est clair de (4.154) que $\partial_y G_J(\theta_J, 0) = 0$. Puisque $F''_{\theta_J}(\theta_J) = 0$, on conclut donc que $\partial_y G_J(x, 0) = \frac{1}{2} F''_{\theta_J}(x)$ pour tout $x \in \mathbb{R}$, ce qui montre *v*). ■

L'inconvénient principal du théorème 4.10 est que les suppositions tenues sur F_{θ_J} et g_J ne sont pas très évocatrices. Heureusement, ces conditions ne sont pas difficiles à vérifier pour des lois de probabilités données. On en donne deux exemples aux sous-sections suivantes.

§ 4.5.1.3. Application à la loi normale

On suppose que les seuils de la population J suivent une loi normale de moyenne θ_J et d'écart type σ_{θ_J} . Alors

$$F'_{\theta_J}(x) = \frac{1}{\sigma_{\theta_J} \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - \theta_J}{\sigma_{\theta_J}}\right)^2\right). \quad (4.156)$$

Dans ce cas, F_{θ_J} est infiniment dérivable d'après les propriétés de la fonction exponentielle. De plus, la loi normale est effectivement unimodale et symétrique. Les conclusions des lemmes 4.8 et 4.9 sont donc respectées pour la loi normale.

On peut alors trouver une expression plus explicite pour g_J . Un calcul direct à partir de l'expression (4.156) montre que

$$F''_{\theta_J}(x) = \frac{\theta_J - x}{\sigma_{\theta_J}^3 \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - \theta_J}{\sigma_{\theta_J}}\right)^2\right) = \frac{\theta_J - x}{\sigma_{\theta_J}^2} F'_{\theta_J}(x). \quad (4.157)$$

Ainsi, en définissant g_J comme au lemme 4.8, on a pour $x \neq \theta_J$ que

$$g_J(x, y) = \frac{y}{2\sigma_{\theta_J}^2}. \quad (4.158)$$

On pourrait calculer la dérivée troisième de F_{θ_J} pour calculer directement $g_J(\theta_J, y)$, mais l'expression précédente correspond à une fonction continue partout sur \mathbb{R}^2 . Comme on sait que g_J est continue sur ce même domaine, (4.158) doit donc être vraie partout.

Ensuite, puisque g_J est constante selon son premier argument, il est clair que pour tout $y \geq 0$ la fonction $x \mapsto g_J(x, y)$ est bornée sur \mathbb{R} , et que $\partial_x g_J \equiv 0$. Comme $g_J \geq 0$, on peut donc conclure que $1 + g_J(x, y) + (\theta_J - x) \partial_x g_J(x, y) \geq 0$ pour tout $(x, y) \in \mathbb{R} \times [0, \infty)$.

Ainsi, avec cette loi de probabilité, toutes les hypothèses du théorème 4.10 sont respectées, et G_J possède les propriétés *i*) à *v*) données au théorème.

§ 4.5.1.4. Application à la loi logistique

On suppose que les seuils de la population J suivent une loi logistique de moyenne θ_J et de facteur d'échelle s_{θ_J} . Alors

$$F_{\theta_J}(x) = \sigma\left(\frac{x - \theta_J}{s_{\theta_J}}\right) \quad \text{où} \quad \sigma(\xi) := \frac{1}{1 + e^{-\xi}}. \quad (4.159)$$

À nouveau, F_{θ_J} est infiniment dérivable d'après les propriétés de la fonction exponentielle. De plus, la loi logistique est unimodale et symétrique. Les conclusions des lemmes 4.8 et 4.9 sont donc vraies pour la loi logistique.

La fonction logistique σ possède la propriété que $\sigma' = \sigma(1 - \sigma)$. À partir de celle-ci, on vérifie aisément que

$$F'_{\theta_J} = \frac{1}{s_{\theta_J}} F_{\theta_J}(1 - F_{\theta_J}), \quad (4.160a)$$

$$F''_{\theta_J} = \frac{1}{s_{\theta_J}^2} F_{\theta_J}(1 - F_{\theta_J})(1 - 2F_{\theta_J}) = \frac{1}{s_{\theta_J}} F'_{\theta_J}(1 - 2F_{\theta_J}), \quad (4.160b)$$

$$F'''_{\theta_J} = \frac{1}{s_{\theta_J}^3} F_{\theta_J}(1 - F_{\theta_J})(1 - 6F_{\theta_J} + 6F_{\theta_J}^2) = \frac{1}{s_{\theta_J}^2} F'_{\theta_J}(1 - 6F_{\theta_J} + 6F_{\theta_J}^2). \quad (4.160c)$$

Ainsi, on voit de la définition de g_J donnée au lemme 4.8 que pour $x \neq \theta_J$,

$$g_J(x, y) = \frac{y}{2s_{\theta_J}} \frac{1 - 2F_{\theta_J}(x)}{\theta_J - x}. \quad (4.161)$$

De plus, puisque $F_{\theta_J}(\theta_J) = 1/2$,

$$F'''_{\theta_J}(\theta_J) = \frac{1}{s_{\theta_J}^2} F'_{\theta_J}(\theta_J)(1 - 3 + 6(1/2)^2) = -\frac{F'_{\theta_J}(\theta_J)}{2s_{\theta_J}^2}, \quad (4.162)$$

de sorte que

$$g_J(\theta_J, y) = \frac{y}{4s_{\theta_J}^2}. \quad (4.163)$$

Pour pouvoir appliquer le théorème 4.10, il reste à montrer que la fonction $x \mapsto g_J(x, y)$ est bornée pour chaque $y \geq 0$, et que $1 + g_J(x, y) + (\theta_J - x) \partial_x g_J(x, y) \geq 0$ pour tout $x \in \mathbb{R}$ et pour tout $y \geq 0$.

On fixe alors $y \geq 0$ et $x \in \mathbb{R}$ tel que $|x - \theta_J| \geq 1$. L'expression (4.161) de g_J montre que

$$|g_J(x, y)| = \frac{y}{2s_{\theta_J}} \left| \frac{1 - 2F_{\theta_J}(x)}{\theta_J - x} \right| \leq \frac{3y}{2s_{\theta_J}}, \quad (4.164)$$

puisque $0 \leq F_{\theta_J} \leq 1$. La fonction $x \mapsto g_J(x, y)$ est donc bornée sur $(-\infty, \theta_J - 1] \cup [\theta_J + 1, \infty)$. Or, cette fonction est continue sur \mathbb{R} , donc en particulier sur $[\theta_J - 1, \theta_J + 1]$. Elle y est donc bornée. On conclut donc qu'elle est bornée partout sur \mathbb{R} .

D'après l'expression donnée pour $\partial_x g_J$ au lemme 4.9, on a pour $x \neq \theta_J$ que

$$(\theta_J - x) \partial_x g_J(x, y) = \frac{y}{2} \left(\frac{1 - 2F_{\theta_J}(x)}{s_{\theta_J}(\theta_J - x)} + \frac{1 - 6F_{\theta_J}(x) + 6F_{\theta_J}^2(x)}{s_{\theta_J}^2} - \frac{(1 - 2F_{\theta_J}(x))^2}{s_{\theta_J}^2} \right) \quad (4.165)$$

$$= \frac{y}{2s_{\theta_J}} \frac{1 - 2F_{\theta_J}(x)}{\theta_J - x} \left(1 - \frac{2}{s_{\theta_J}} \frac{\theta_J - x}{1 - 2F_{\theta_J}(x)} (F_{\theta_J}(x) - F_{\theta_J}^2(x)) \right). \quad (4.166)$$

En comparant avec (4.161), on trouve que

$$(\theta_J - x) \partial_x g_J(x, y) = g_J(x, y) \left(1 + 2 \frac{x - \theta_J}{s_{\theta_J}} \frac{F_{\theta_J}(x)(1 - F_{\theta_J}(x))}{1 - 2F_{\theta_J}(x)} \right). \quad (4.167)$$

Pour retravailler cette expression, on remarque d'après l'expression (4.159) de F_{θ_J} , en utilisant à nouveau $\xi := (x - \theta_J)/s_{\theta_J}$, que

$$1 - F_{\theta_J}(x) = \frac{e^{-\xi}}{1 + e^{-\xi}} \quad \text{et que} \quad 1 - 2F_{\theta_J}(x) = -\frac{1 - e^{-\xi}}{1 + e^{-\xi}}, \quad (4.168)$$

de sorte que

$$\frac{F_{\theta_J}(x)(1 - F_{\theta_J}(x))}{1 - 2F_{\theta_J}(x)} = -\frac{1}{1 + e^{-\xi}} \frac{e^{-\xi}}{1 + e^{-\xi}} \frac{1 + e^{-\xi}}{1 - e^{-\xi}} = -\frac{e^{-\xi}}{1 - e^{-2\xi}} = -\frac{1}{e^\xi - e^{-\xi}}. \quad (4.169)$$

On trouve ainsi que

$$(\theta_J - x) \partial_x g_J(x, y) = g_J(x, y) \left(1 - \frac{2\xi}{e^\xi - e^{-\xi}} \right) = g_J(x, y) \left(1 - \frac{\xi}{\sinh \xi} \right). \quad (4.170)$$

Or, on sait que $|\xi| \leq |\sinh \xi|$ pour tout $\xi \in \mathbb{R}$, donc que $|\frac{\xi}{\sinh \xi}| \leq 1$ pour tout $\xi \neq 0$. Puisque g_J est positive, on conclut donc que $(\theta_J - x) \partial_x g_J(x, y) \geq 0$ pour tout $x \neq \theta_J$ et pour tout $y \geq 0$. De plus, on sait d'après le lemme 4.9 que $\partial_x g_J(\theta_J, y) = 0$ pour tout $y \geq 0$. On conclut ainsi que l'expression $1 + g_J(x, y) + (\theta_J - x) \partial_x g_J(x, y)$ est positive pour tout $x \in \mathbb{R}$ et pour tout $y \geq 0$.

Ainsi, avec cette loi de probabilité, toutes les hypothèses du théorème 4.10 sont respectées, et G_J possède les propriétés i) à v) données au théorème.

§ 4.5.2. Fermeture de moments

À partir de la fonction G_J étudiée à la sous-section précédente, on peut donner des approximations des espérances de produits qui impliquent des fonctions de répartition, en vérifiant à chaque fois que ces approximations sont cohérentes avec les approximations de champ moyen de premier et de deuxième ordre.

Les espérances dont on veut trouver des approximations pour rendre le système (4.16)–(4.42) autonome sont de formes

$$\mathbb{E}[S_t^J F_{\theta_j}(B_t^J)], \quad \mathbb{E}[A_t^J S_t^K F_{\theta_K}(B_t^K)] \quad \text{et} \quad \mathbb{E}[R_t^J S_t^K F_{\theta_K}(B_t^K)] \quad (4.171)$$

pour des populations J et K du graphe. On traite ces deux cas séparément.

D'abord, on cherche à approximer $\mathbb{E}[S_t^J F_{\theta_j}(B_t^J)]$ par une fonction des espérances de S_t^J et de B_t^J , de leur covariance, ainsi que de la variance de B_t^J puisqu'elle pourrait intervenir de la même façon que dans l'approximation de $\mathbb{E}[F_{\theta_j}(B_t^J)]$ construite à la section précédente. On peut obtenir plusieurs conditions que doit posséder une telle approximation. D'abord, si les variances de S_t^J et de B_t^J sont nulles, on sait que $\mathbb{E}[S_t^J F_{\theta_j}(B_t^J)] = S_t^J(t) F_{\theta_j}(\mathcal{B}_J(t))$. De même, si la covariance entre S_t^J et B_t^J est nulle, on s'attend à ce que $\mathbb{E}[S_t^J F_{\theta_j}(B_t^J)] = S_J(t) \mathbb{E}[F_{\theta_j}(B_t^J)]$. Ceci donne une condition claire si l'on convient que l'on approxime $\mathbb{E}[F_{\theta_j}(B_t^J)] \approx G_J(\mathcal{B}_J(t), C_{BB}^{JJ}(t))$. Ensuite, puisque F_{θ_j} est croissante, on sait que le signe de la covariance $\text{Cov}[S_t^J, F_{\theta_j}(B_t^J)] = \mathbb{E}[S_t^J F_{\theta_j}(B_t^J)] - S_J(t) \mathbb{E}[F_{\theta_j}(B_t^J)]$ devrait suivre le signe de la covariance entre S_t^J et B_t^J .

Des conditions sur les comportements limites de $\mathbb{E}[S_t^J F_{\theta_j}(B_t^J)]$ peuvent aussi être obtenues. En effet, si $\mathcal{B}_J(t)$ est très grande positivement, $F_{\theta_j}(B_t^J)$ est très près de 1 avec une probabilité très élevée. On s'attend alors à ce que $\mathbb{E}[S_t^J F_{\theta_j}(B_t^J)]$ soit près de $\mathcal{B}_J(t)$. De la même façon, si $\mathcal{B}_J(t)$ est très grande négativement, $F_{\theta_j}(B_t^J)$ est très près de zéro avec une probabilité élevée et l'on s'attend à ce que $\mathbb{E}[S_t^J F_{\theta_j}(B_t^J)]$ soit aussi près de zéro.

Finalement, on peut trouver quelques conditions pour assurer que l'approximation obtenue soit cohérente avec l'approximation de Taylor d'ordre 2. En effet, on a vu en (4.46) l'approximation

$$\mathbb{E}[S_t^J F_{\theta_j}(B_t^J)] \approx S_J(t) F_{\theta_j}(\mathcal{B}_J(t)) + F'_{\theta_j}(\mathcal{B}_J(t)) C_{SB}^{JJ}(t) + \frac{1}{2} S_J(t) F''_{\theta_j}(\mathcal{B}_J(t)) C_{BB}^{JJ}(t). \quad (4.46)$$

Celle-ci devrait toujours être bonne pour de petites covariances, ce qui impose des conditions sur les dérivées de l'approximation évaluées à des covariances nulles.

Le résultat suivant introduit une fonction qui remplit les conditions énumérées plus haut et qui est suffisamment régulière, du moins sur un domaine restreint. Pour l'énoncé suivant, on omet les indices relatifs aux populations sur les variables puisque leur signification n'est pas pertinente, mais on utilise tout de même les mêmes lettres afin de faciliter le lien avec les explications précédentes.

Proposition 4.11. Soient F_{θ_J} , g_J et G_J les fonctions introduites au lemme 4.8 et au théorème 4.10. Soit $f: (0, \infty) \times \mathbb{R}^3 \rightarrow [0, 1]$ donnée par

$$f(\mathcal{S}, \mathcal{B}, C_{SB}, C_{BB}) := \mathcal{S}G_J\left(\mathcal{B} + \frac{C_{SB}}{\mathcal{S}}, C_{BB}\right). \quad (4.172)$$

Alors :

- i) f est de classe \mathcal{C}^1 sur $(0, \infty) \times \mathbb{R}^3$;
- ii) Pour $\mathcal{S} > 0$ et $\mathcal{B} \in \mathbb{R}$, $f(\mathcal{S}, \mathcal{B}, 0, 0) = \mathcal{S}F_{\theta_J}(\mathcal{B})$;
- iii) Pour $\mathcal{S} > 0$, $\mathcal{B}, C_{SB} \in \mathbb{R}$ et $C_{BB} \geq 0$, $f(\mathcal{S}, \mathcal{B}, C_{SB}, C_{BB}) \geq \mathcal{S}G_J(\mathcal{B}, C_{BB})$ si $C_{SB} \geq 0$;
- iv) Pour $\mathcal{S} > 0$, $\mathcal{B}, C_{SB} \in \mathbb{R}$ et $C_{BB} \geq 0$, $f(\mathcal{S}, \mathcal{B}, C_{SB}, C_{BB}) \rightarrow \mathcal{S}$ lorsque $\mathcal{B} \rightarrow \infty$ et $f(\mathcal{S}, \mathcal{B}, C_{SB}, C_{BB}) \rightarrow 0$ lorsque $\mathcal{B} \rightarrow -\infty$;
- v) Pour $\mathcal{S} > 0$ et $\mathcal{B} \in \mathbb{R}$, $\partial_{C_{SB}} f(\mathcal{S}, \mathcal{B}, 0, 0) = F'_{\theta_J}(\mathcal{B})$;
- vi) Pour $\mathcal{S} > 0$ et $\mathcal{B} \in \mathbb{R}$, $\partial_{C_{BB}^{II}} f(\mathcal{S}, \mathcal{B}, 0, 0) = \frac{1}{2} \mathcal{S}F''_{\theta_J}(\mathcal{B})$. ◆

Démonstration. Ces propriétés sont toutes des conséquences assez directes des propriétés de G_J démontrées au théorème 4.10. On démontre ces résultats point par point.

- i) On sait que G_J est de classe \mathcal{C}^1 sur \mathbb{R}^2 . Puisque la fonction $(\mathcal{S}, \mathcal{B}, C_{SB}) \mapsto \mathcal{B} + C_{SB}/\mathcal{S}$ est de classe \mathcal{C}^1 sur $(0, \infty) \times \mathbb{R}^2$, alors f est de classe \mathcal{C}^1 sur $(0, \infty) \times \mathbb{R}^3$.
- ii) On sait que $G_J(x, 0) = F_{\theta_J}(x)$ pour tout $x \in \mathbb{R}$. Ainsi, $f(\mathcal{S}, \mathcal{B}, 0, 0) = \mathcal{S}G_J(\mathcal{B}, 0) = \mathcal{S}F_{\theta_J}(\mathcal{B})$ pour $\mathcal{S} > 0$ et pour $\mathcal{B} \in \mathbb{R}$.
- iii) On a vu que la fonction $x \mapsto G_J(x, y)$ est croissante sur \mathbb{R} pour tout $y \geq 0$. On déduit donc que pour $\mathcal{S} > 0$, pour $\mathcal{B}, C_{SB} \in \mathbb{R}$ et pour $C_{BB} \geq 0$,

$$f(\mathcal{S}, \mathcal{B}, C_{SB}, C_{BB}) = \mathcal{S}\left(\mathcal{B} + \frac{C_{SB}}{\mathcal{S}}, C_{BB}\right) \geq \mathcal{S}G_J(\mathcal{B}, C_{BB}) \quad \text{si } C_{SB} \geq 0. \quad (4.173)$$

- iv) On a vu que si $x \in \mathbb{R}$ et $y \geq 0$, alors $G_J(x, y) \rightarrow 0$ lorsque $x \rightarrow -\infty$ et $G_J(x, y) \rightarrow 1$ lorsque $x \rightarrow \infty$. Pour $\mathcal{S} > 0$, pour $\mathcal{B}, C_{SB} \in \mathbb{R}$ et pour $C_{BB} \geq 0$, on voit alors que $f(\mathcal{S}, \mathcal{B}, C_{SB}, C_{BB}) = \mathcal{S}G_J(\mathcal{B} + C_{SB}/\mathcal{S}, C_{BB}) \rightarrow 0$ lorsque $\mathcal{B} \rightarrow -\infty$, et $f(\mathcal{S}, \mathcal{B}, C_{SB}, C_{BB}) = \mathcal{S}G_J(\mathcal{B} + C_{SB}/\mathcal{S}, C_{BB}) \rightarrow \mathcal{S}$ lorsque $\mathcal{B} \rightarrow \infty$.
- v) D'après l'expression (4.152) de la dérivée de G_J selon sa première composante, puisque $g_J(x, 0) = \partial_x g_J(x, 0) = 0$ pour tout $x \in \mathbb{R}$, on voit directement que si $\mathcal{S} > 0$ et $\mathcal{B} \in \mathbb{R}$,

$$\partial_{C_{SB}} f(\mathcal{S}, \mathcal{B}, 0, 0) = \mathcal{S}F'_{\theta_J}(\mathcal{B}) \cdot \frac{1}{\mathcal{S}} = F'_{\theta_J}(\mathcal{B}). \quad (4.174)$$

- vi) On a vu au théorème 4.10 que $\partial_y G_J(x, 0) = \frac{1}{2} F''_{\theta_J}(x)$ pour tout $x \in \mathbb{R}$. Ainsi, si $\mathcal{S} > 0$ et $\mathcal{B} \in \mathbb{R}$, alors $\partial_{C_{BB}^{II}} f(\mathcal{S}, \mathcal{B}, 0, 0) = \frac{1}{2} \mathcal{S}F''_{\theta_J}(\mathcal{B})$. ■

On cherche maintenant à trouver une approximation de $\mathbb{E}[A_t^J S_t^K F_{\theta_K}(B_t^K)]$ par une fonction des espérances de A_t^J , de S_t^K et de B_t^K ainsi que de leurs covariances et de la variance de B_t^K . On cherche du même coup une approximation des espérances de la forme $\mathbb{E}[R_t^J S_t^K F_{\theta_K}(B_t^K)]$, puisque les deux formes sont très similaires. De la même façon que dans le dernier cas, on peut obtenir quelques conditions sur l'approximation voulue.

D'abord, si les variances de ces variables aléatoires sont nulles, leurs covariances le sont toutes aussi et $\mathbb{E}[A_t^J S_t^K F_{\theta_K}(B_t^K)] = \mathcal{A}_J(t) \mathcal{S}_K(t) F_{\theta_K}(B_K(t))$. De même, si elles ne sont pas corrélées, alors $\mathbb{E}[A_t^J S_t^K F_{\theta_K}(B_t^K)] = \mathcal{A}_J(t) \mathcal{S}_K(t) \mathbb{E}[F_{\theta_K}(B_t^K)]$ ce qui, à nouveau, donne une condition claire en utilisant la fonction G_J pour approximer l'espérance de $F_{\theta_K}(B_t^K)$.

Ensuite, on trouve quelques conditions pour assurer que l'approximation obtenue soit cohérente avec l'approximation de Taylor d'ordre 2. En effet, on a vu en (4.52) l'approximation

$$\text{Cov}[A_t^J, S_t^K F_{\theta_K}(B_t^K)] \approx F_{\theta_K}(B_K(t)) C_{AS}^{JK}(t) + \mathcal{S}_K(t) F'_{\theta_K}(B_K(t)) C_{AB}^{JK}(t). \quad (4.52)$$

En lui additionnant le produit de $\mathcal{A}_J(t)$ et de l'approximation (4.46) pour l'espérance de $S_t^K F_{\theta_K}(B_t^K)$, on obtient que

$$\mathbb{E}[A_t^J S_t^K F_{\theta_K}(B_t^K)] = \mathcal{A}_J(t) \mathbb{E}[S_t^K F_{\theta_K}(B_t^K)] + \text{Cov}[A_t^J, S_t^K F_{\theta_K}(B_t^K)] \quad (4.175)$$

$$\begin{aligned} &\approx \mathcal{A}_J(t) \left(\mathcal{S}_K(t) F_{\theta_K}(B_K(t)) + F'_{\theta_K}(B_K(t)) C_{SB}^{KK}(t) \right. \\ &\quad \left. + \frac{1}{2} \mathcal{S}_K(t) F''_{\theta_K}(B_K(t)) C_{BB}^{KK}(t) \right) \end{aligned} \quad (4.176)$$

$$\begin{aligned} &+ F_{\theta_K}(B_K(t)) C_{AS}^{JK}(t) + \mathcal{S}_K(t) F'_{\theta_K}(B_K(t)) C_{AB}^{JK}(t) \\ &= (\mathcal{A}_J(t) \mathcal{S}_K(t) + C_{AS}^{JK}(t)) F_{\theta_K}(B_K(t)) \\ &\quad + (\mathcal{A}_J(t) C_{SB}^{KK}(t) + \mathcal{S}_K(t) C_{AB}^{JK}(t)) F'_{\theta_K}(B_K(t)) \\ &\quad + \frac{1}{2} \mathcal{A}_J(t) \mathcal{S}_K(t) F''_{\theta_K}(B_K(t)) C_{BB}^{KK}(t). \end{aligned} \quad (4.177)$$

Comme cette approximation devrait être bonne pour de petites covariances, elle impose des conditions sur les dérivées de l'approximation recherchée lorsque évaluée à des covariances nulles.

Le résultat suivant introduit une fonction qui remplit les conditions énumérées plus haut et qui est suffisamment régulière, mais seulement sur un domaine restreint, de la même façon que dans le dernier cas.

Proposition 4.12. Soient F_{θ_J} , g_J et G_J les fonctions introduites au lemme 4.8 et au théorème 4.10. Soit $f: (0, \infty)^2 \times \mathbb{R}^5 \rightarrow [0, 1]$ donnée par

$$f(\mathcal{A}, \mathcal{S}, \mathcal{B}, C_{AS}, C_{AB}, C_{SB}, C_{BB}) := (\mathcal{A}\mathcal{S} + C_{AS}) G_J \left(\mathcal{B} + \frac{C_{AB}}{\mathcal{A}} + \frac{C_{SB}}{\mathcal{S}}, C_{BB} \right). \quad (4.178)$$

Alors :

- i) f est de classe \mathcal{C}^1 sur $(0, \infty)^2 \times \mathbb{R}^5$;
- ii) Pour $\mathcal{A}, \mathcal{S} > 0$, $\mathcal{B} \in \mathbb{R}$ et $C_{BB} \geq 0$, alors $f(\mathcal{S}, \mathcal{B}, 0, 0, 0, C_{BB}) = \mathcal{A}\mathcal{S}G_J(\mathcal{B}, C_{BB})$;
- iii) Pour $\mathcal{A}, \mathcal{S} > 0$ et $\mathcal{B} \in \mathbb{R}$, $f(\mathcal{A}, \mathcal{S}, \mathcal{B}, 0, 0, 0, 0) = \mathcal{A}\mathcal{S}F_{\theta_J}(\mathcal{B})$;
- iv) Pour $\mathcal{A}, \mathcal{S} > 0$ et $\mathcal{B} \in \mathbb{R}$, $\partial_{C_{AS}} f(\mathcal{A}, \mathcal{S}, \mathcal{B}, 0, 0, 0, 0) = F_{\theta_J}(\mathcal{B})$;
- v) Pour $\mathcal{A}, \mathcal{S} > 0$ et $\mathcal{B} \in \mathbb{R}$, $\partial_{C_{AB}} f(\mathcal{A}, \mathcal{S}, \mathcal{B}, 0, 0, 0, 0) = \mathcal{S}F'_{\theta_J}(\mathcal{B})$;
- vi) Pour $\mathcal{A}, \mathcal{S} > 0$ et $\mathcal{B} \in \mathbb{R}$, $\partial_{C_{SB}} f(\mathcal{A}, \mathcal{S}, \mathcal{B}, 0, 0, 0, 0) = \mathcal{A}F'_{\theta_J}(\mathcal{B})$;
- vii) Pour $\mathcal{A}, \mathcal{S} > 0$ et $\mathcal{B} \in \mathbb{R}$, $\partial_{C_{BB}} f(\mathcal{A}, \mathcal{S}, \mathcal{B}, 0, 0, 0, 0) = \frac{1}{2} \mathcal{A}\mathcal{S}F''_{\theta_J}(\mathcal{B})$.

◆

Remarque. Ce résultat donne une approximation de $\mathbb{E}[A_t^J S_t^K F_{\theta_K}(B_t^K)]$ qui respecte les critères énoncés plus tôt, mais montre aussi que l'approximation

$$\mathbb{E}[R_t^J S_t^K F_{\theta_K}(B_t^K)] \approx (\mathcal{R}_J(t) \mathcal{S}_K(t) + C_{RS}^{JK}(t)) G_K \left(\mathcal{B}_K(t) + \frac{C_{RB}^{JK}(t)}{\mathcal{R}_J(t)} + \frac{C_{SB}^{KK}(t)}{\mathcal{S}_K(t)}, C_{BB}^{KK}(t) \right) \quad (4.179)$$

respecte les critères analogues.

Démonstration. À nouveau, ces résultats sont tous des conséquences assez directes des propriétés de G_J démontrées au théorème 4.10. On abrège donc les détails.

D'abord, le résultat *i*) suit du fait que G_J est de classe \mathcal{C}^1 sur \mathbb{R}^2 , puisque les cas où $\mathcal{A} = 0$ ou $\mathcal{S} = 0$ sont exclus du domaine de f . Les résultats *ii*) et *iii*) suivent ensuite directement de l'évaluation de f , puisque $G_J(x, 0) = F_{\theta_J}(x)$ pour tout $x \in \mathbb{R}$. Le résultat *iv*) suit de la même propriété de G_J une fois la dérivée par rapport à C_{AS} calculée.

Ensuite, on a vu à partir de l'expression (4.152) de $\partial_x G_J$ que $\partial_x G_J(x, 0) = F'_{\theta_J}(x)$ pour tout $x \in \mathbb{R}$. Les résultats *v*) et *vi*) suivent alors de l'évaluation de $\partial_{C_{AB}} f$ et de $\partial_{C_{SB}} f$. Finalement, on a vu que $\partial_y G_J(x, 0) = \frac{1}{2} F''_{\theta_J}(x)$ pour tout $x \in \mathbb{R}$, ce qui implique le résultat *vii*). ■

Les fonctions introduites aux deux propositions précédentes suffisent à donner une solution au problème de fermeture de moments pour le système (4.16)–(4.42). Pour l'écrire de façon plus concise, on introduit des fonctions $H_J : (0, \infty)^2 \times \mathbb{R}^5 \rightarrow \mathbb{R}$ données par

$$H_J(\mathcal{A}, \mathcal{S}, \mathcal{B}, C_{AS}, C_{AB}, C_{SB}, C_{BB}) := (\mathcal{A}\mathcal{S} + C_{AS})G_J \left(\mathcal{B} + \frac{C_{AB}}{\mathcal{A}} + \frac{C_{SB}}{\mathcal{S}}, C_{BB} \right) - \mathcal{A}\mathcal{S}G_J \left(\mathcal{B} + \frac{C_{SB}}{\mathcal{S}}, C_{BB} \right). \quad (4.180)$$

Ces fonctions permettent de représenter les approximations des covariances de formes $\text{Cov}[A_t^J S_t^K F_{\theta_K}(B_t^K)]$ et $\text{Cov}[R_t^J S_t^K F_{\theta_K}(B_t^K)]$ prévues par les propositions 4.11 et 4.12.

Ces approximations mènent alors au système, pour des populations J et K du graphe,

$$\dot{\mathcal{A}}_J = -\beta_J \mathcal{A}_J + \alpha_J \mathcal{S}_J G_J \left(\mathcal{B}_J + \frac{C_{SB}^{JJ}}{\mathcal{S}_J}, C_{BB}^{JJ} \right), \quad (4.181a)$$

$$\dot{\mathcal{R}}_J = -\gamma_J \mathcal{R}_J + \beta_J \mathcal{A}_J, \quad (4.181b)$$

$$\begin{aligned} \dot{C}_{AA}^{JK} &= -(\beta_J + \beta_K) C_{AA}^{JK} + \alpha_K H_K(\mathcal{A}_J, \mathcal{S}_K, \mathcal{B}_K, C_{AS}^{JK}, C_{AB}^{JK}, C_{SB}^{KK}, C_{BB}^{KK}) \\ &\quad + \alpha_J H_J(\mathcal{A}_K, \mathcal{S}_J, \mathcal{B}_J, C_{AS}^{KJ}, C_{AB}^{KJ}, C_{SB}^{JJ}, C_{BB}^{JJ}), \end{aligned} \quad (4.181c)$$

$$\dot{C}_{RR}^{JK} = -(\gamma_J + \gamma_K) C_{RR}^{JK} + \beta_K C_{AR}^{KJ} + \beta_J C_{AR}^{JK}, \quad (4.181d)$$

$$\dot{C}_{AR}^{JK} = -(\beta_J + \gamma_K) C_{AR}^{JK} + \beta_K C_{AA}^{JK} + \alpha_J H_J(\mathcal{R}_K, \mathcal{S}_J, \mathcal{B}_J, C_{RS}^{KJ}, C_{RB}^{KJ}, C_{SB}^{JJ}, C_{BB}^{JJ}). \quad (4.181e)$$

Comme dans le cas du système (4.54) obtenu de l'approximation de champ moyen d'ordre 2, toutes les variables utilisées dans ce système peuvent être exprimées seulement en fonction des fractions actives et réfractaires des populations du graphe. Ainsi, le système (4.181) est

bien autonome. De plus, ce système constitue une extension du modèle de champ moyen (4.17) dans le même sens que le système (4.54). En effet, si l'on néglige toutes les covariances dans le système (4.181), d'après les propriétés des fonctions G_J démontrées au théorème 4.10 on retrouve exactement le système moyen (4.17).

Finalement, il est bon de remarquer que même si ce système est défini dans $\mathbb{R}^{n(2n+3)}$ de la même façon que le système (4.54), ce n'est pas exactement sur ce domaine que l'on peut facilement étudier ses solutions. En effet, on a montré aux propositions 4.11 et 4.12 que si les F_{θ_J} sont de classe \mathcal{C}^4 , alors les fonctions utilisées pour approximer les espérances de produits sont de classe \mathcal{C}^1 , mais seulement sur des domaines restreints : on doit exclure les points où une composante de la forme \mathcal{A}_J , \mathcal{R}_J ou \mathcal{S}_J devient nulle. Dans ce cas, le théorème 1.2 montre que le système (4.181) a des solutions uniques sur $\mathcal{D}_n^\circ \times \mathbb{R}^{n(2n+1)}$, où

$$\mathcal{D}_n^\circ = \{(\mathcal{A}_J, \mathcal{R}_J)_{J \in \mathcal{P}} \in (0, 1)^{2n} : \forall J \in \mathcal{P}, \mathcal{A}_J + \mathcal{R}_J < 1\} \quad (4.182)$$

est l'intérieur du domaine \mathcal{D}_n défini en (4.23).

Étant donnée la forme des équations, ce système est encore plus difficile à étudier analytiquement que le système (4.54) obtenu de l'approximation de champ moyen de deuxième ordre. L'analyse qu'on en a faite jusqu'ici est seulement basée sur des exemples numériques.

§ 4.5.3. Quelques résultats numériques

On présente ici quelques exemples où le système (4.181) est étudié numériquement. On revient d'abord sur les exemples présentés aux sections 4.4.2.4 et 4.4.3, où le système obtenu de l'approximation de champ moyen de deuxième ordre prévoit des résultats qui n'ont pas de sens ou ne concordent pas avec les simulations du processus stochastique. On se concentre ensuite sur l'étude d'un nouvel exemple où le modèle (4.181) et le modèle de champ moyen mènent à des prédictions différentes, et où la prédition du modèle (4.181) est plus près des trajectoires prévues par les simulations du processus stochastique.

§ 4.5.3.1. Comparaison avec le système de champ moyen de deuxième ordre

On a vu à la section 4.4.2.4 qu'une des faiblesses importantes du système (4.54) issu de l'approximation de champ moyen de deuxième ordre est qu'il prévoit parfois des solutions dont les composantes reliées aux covariances sortent du domaine physiologique (4.67), qui est le domaine où les valeurs de ces variables sont cohérentes avec l'interprétation qu'on veut en faire. Ce problème est bien illustré par l'exemple pour lequel les solutions sont présentées à la figure 4.10.

Parmi les solutions du système (4.181) obtenues par des intégrations numériques, on n'a pas vu de cas où les composantes associées aux covariances se mettent à grandir de façon exponentielle de la façon illustrée à la figure 4.10. Par exemple, avec l'état initial (4.122) et

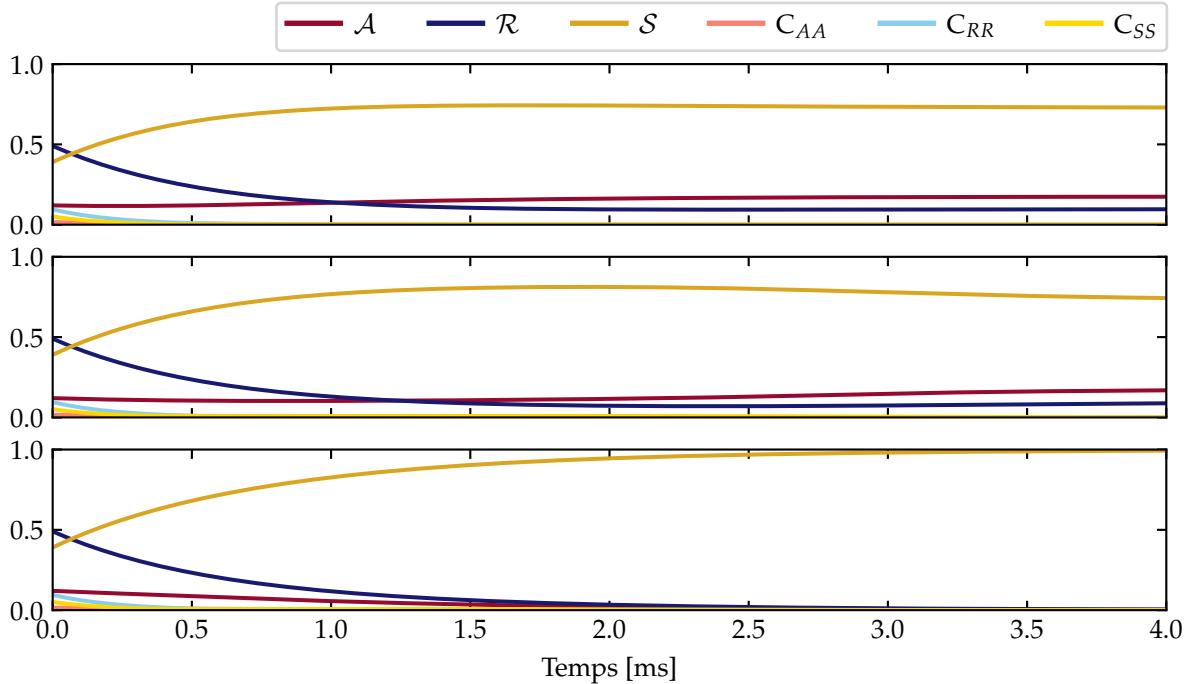


FIGURE 4.16 – Solutions du système (4.181) pour l'état initial (4.122) et pour les paramètres (4.121), avec $Q = 0$ en haut, $Q = -1/2$ au milieu et $Q = -3/2$ en bas.

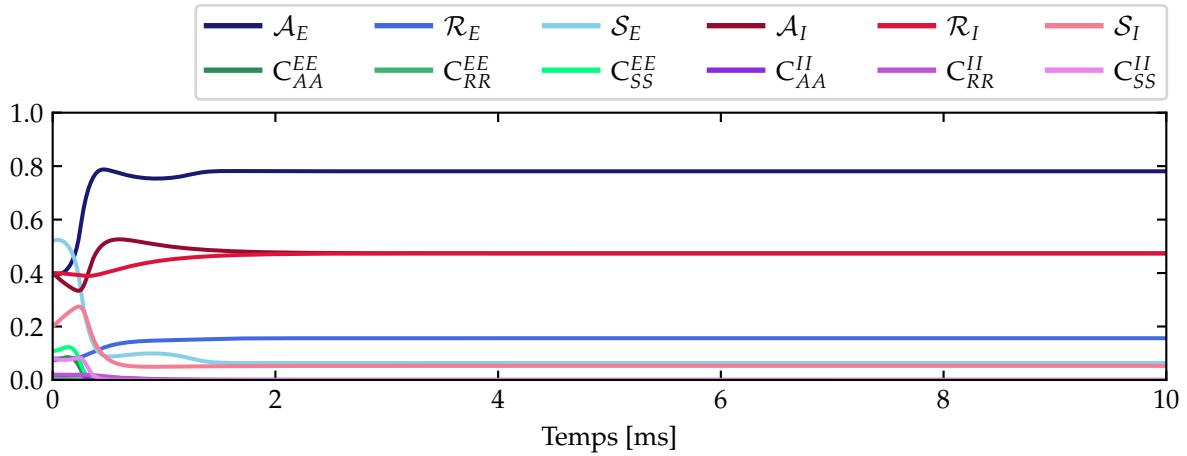


FIGURE 4.17 – Solution du système (4.181) pour les paramètres (4.123) et l'état initial (4.124).

les paramètres (4.121) utilisés pour obtenir les solutions de la figure 4.10, le système (4.181) a les solutions illustrées à la figure 4.16 pour les trois valeurs d'activité entrante externe. Dans les trois cas, les variances diminuent rapidement vers zéro et les solutions concordent toutes avec les trajectoires du processus stochastiques présentées à la figure 4.11.

On a aussi vu à la section 4.4.3 un exemple où le système (4.54) obtenu de l'approximation de champ moyen de deuxième ordre menait à une solution très différente de celle prévue par le modèle de champ moyen, mais où les trajectoires du processus stochastique prévues par les simulations numériques correspondaient davantage aux solutions du modèle de

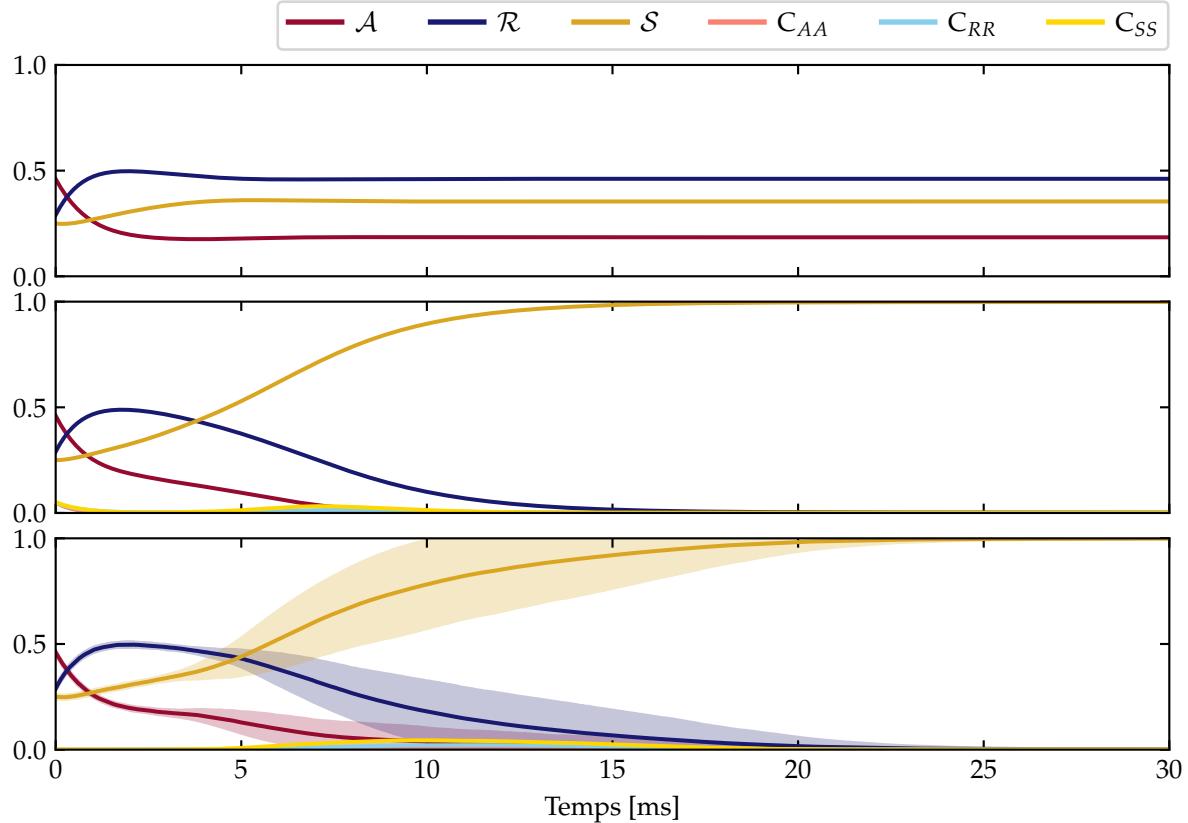


FIGURE 4.18 – En haut, solution du système (4.17) obtenu de l’approximation de champ moyen pour les paramètres (4.183) et l’état initial (4.184) où l’on ignore les covariances. Au milieu, solution du système (4.181) pour les mêmes paramètres et le même état initial, sans ignorer les covariances. En bas, statistiques obtenues à partir de 100 trajectoires possibles du processus stochastique. Les régions colorées autour des courbes qui représentent les valeurs moyennes sont bornées par un écart type de chaque côté de la moyenne. Voir l’annexe B.2 pour les détails sur la méthodologie employée pour les simulations.

champ moyen. La solution du système (4.181) obtenue numériquement, qui est présentée à la figure 4.17, concorde avec celle du système de champ moyen ainsi qu’avec les trajectoires du processus stochastique.

Ces résultats suggèrent que cette nouvelle extension du modèle de champ moyen ne souffre pas des faiblesses du modèle de champ moyen de deuxième ordre (4.54).

§ 4.5.3.2. Étude d’un exemple à une population

On présente ici un exemple où la solution du système (4.181) converge vers un état stable distinct de celui vers lequel la solution du système de champ moyen (4.17) converge.

On considère un graphe constitué d’une seule population. On suppose que les seuils suivent une loi logistique avec une fonction de répartition de la forme (4.31a), et on fixe les para-

mètres

$$\begin{aligned} \alpha &= 0,56 \text{ ms}^{-1}, & \beta &= 1 \text{ ms}^{-1}, & \gamma &= 0,4 \text{ ms}^{-1}, \\ \theta &= 0,73, & s_\theta &= 0,1, & Q &= 0, & c &= 5,36 \end{aligned} \quad (4.183)$$

et l'état initial

$$(\mathcal{A}_0, \mathcal{R}_0, C_{AA0}, C_{RR0}, C_{AR0}) := (0,46; 0,29; 0,05; 0,05; -0,025), \quad (4.184)$$

où l'on ignore simplement les covariances pour obtenir la condition initiale à considérer pour le modèle de champ moyen.

En intégrant numériquement les équations différentielles (4.181) et (4.17) pour ces paramètres, on obtient les solutions présentées à la figure 4.18. En plus des intégrations numériques, 100 trajectoires possibles du processus stochastique sous-jacent ont été obtenues à partir de simulations numériques. À partir de ces trajectoires, on a obtenu une estimation des valeurs moyennes et des variances des fractions de la population dans chacun des trois états. Ces résultats sont aussi illustrés à la figure 4.18.

D'après les résultats de la figure 4.18, il est clair que les résultats prévus par le modèle (4.181), qui inclut les covariances, correspondent mieux aux résultats prévus par les trajectoires du processus stochastique que ceux prévus par le modèle de champ moyen.

Pour mieux comprendre ce qui se produit, il est instructif de tracer un portrait de phase du système de champ moyen (4.17) pour ces paramètres. On trace également sur ce portrait de phase la solution du système de champ moyen, ainsi que la trajectoire donnée par les simulations du processus stochastique. Le tout est illustré à la figure 4.19. On voit bien que les trajectoires du processus stochastique sortent du bassin d'attraction du point fixe vers lequel converge la solution du système de champ moyen.

Puisque le système (4.181) qui inclut les covariances a ici cinq dimensions, on ne peut pas tracer son portrait de phase. Cependant, on peut toujours tracer sa projection

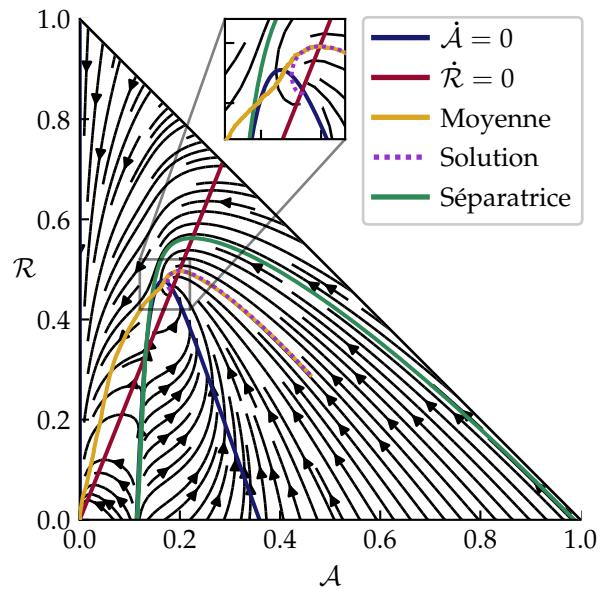


FIGURE 4.19 – Portrait de phase du système de champ moyen (4.17) pour les paramètres (4.183). La solution tracée est celle du système de champ moyen, et la moyenne indiquée est la trajectoire obtenue des statistiques illustrées en bas de la figure 4.18. La courbe qui représente la séparatrice entre les bassins d'attraction des deux points fixes stables a été estimée en intégrant à l'envers dans le temps à partir d'états près du point fixe instable.

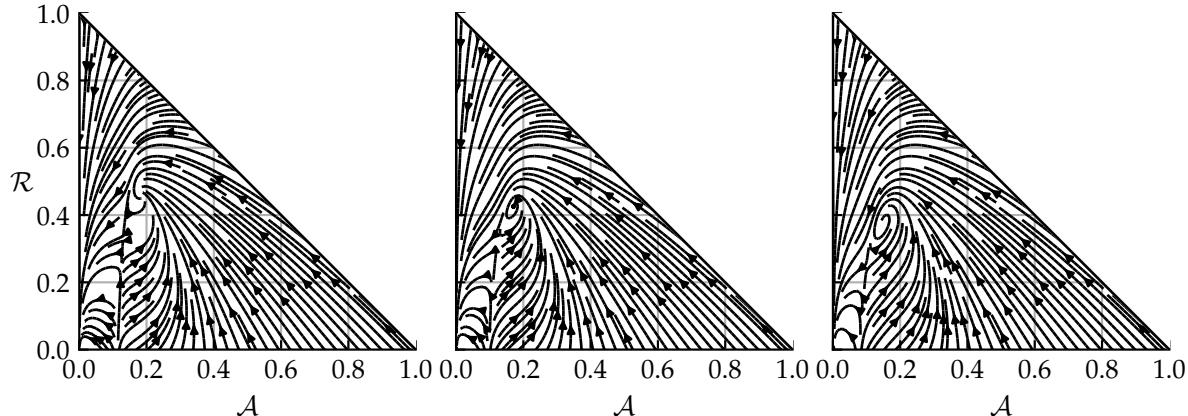


FIGURE 4.20 – Projections des portraits de phase du système (4.181) dans le plan \mathcal{A} - \mathcal{R} pour des variances de 0, 1/1000 et 2/1000.

dans le plan \mathcal{A} - \mathcal{R} en fixant des valeurs pour les covariances. Trois exemples de telles projections sont présentés à la figure 4.20. Il est intéressant d'y observer particulièrement la région du plan de phase pour des valeurs de $(\mathcal{A}, \mathcal{R})$ autour de $(0,2; 0,5)$. Dans cette région, des solutions convergent vers le point fixe à zéro pour des valeurs de \mathcal{R} plus faibles lorsque les variances augmentent. Ceci semble cohérent avec le comportement des trajectoires du processus stochastique, selon la courbe illustrée à la figure 4.19.

Dans ce cas ci, le modèle (4.181) permet donc de reproduire le comportement du processus stochastique, alors que le modèle de champ moyen ne peut pas le faire.

§ 4.5.4. Conclusion

On a proposé ici une nouvelle solution au problème de fermeture de moments pour le système (4.16)–(4.42), en basant les approximations d'espérances de produits qui impliquent des fonctions de forme sigmoïdale sur des observations qualitatives sur la forme de ces espérances. On a ainsi obtenu le système dynamique (4.181).

Ce nouveau système dynamique semble prometteur pour fournir une bonne approximation du comportement macroscopique de la dynamique MTC présenté à la section 4.2 afin de représenter l'activité d'un réseau de neurones biologiques, tout en permettant des corrélations non nulles entre les fractions des populations qui se trouvent dans chacun des trois états. En effet, on a vu plus tôt que le modèle issu de l'approximation de champ moyen de deuxième ordre et présenté à la section 4.4 avait des faiblesses importantes : il prévoit des solutions qui sortent du domaine où elles peuvent avoir un sens par rapport aux variables aléatoires qu'elles servent à représenter, et il prévoit des solutions qui ont un comportement tout autre que celui des trajectoires du processus stochastique. Or, les exemples présentés dans cette section suggèrent que le nouveau système dynamique (4.181) ne souffre pas de ces problèmes.

De plus, on a vu un exemple d'un cas où le système (4.181) prévoit des solutions qui se rapprochent davantage des trajectoires du processus stochastique que les solutions du modèle de champ moyen (4.17). Si cet exemple est encourageant, il serait néanmoins intéressant de trouver d'autres cas où les prévisions du système (4.181) sont plus près du comportement du processus stochastique que celles du modèle de champ moyen, notamment lorsque le graphe est divisé en davantage de populations. En effet, dans ces cas, le système a davantage de dimensions, donc des dynamiques plus variées sont possibles.

Conclusion

La modélisation mathématique de phénomènes biologiques qui impliquent un très grand nombre d'unités est une tâche ardue, d'autant plus lorsque l'on veut décrire des corrélations entre les états de ces unités. Le modèle de Wilson–Cowan n'échappe pas à cette tendance. En effet, la version classique de ce modèle ne décrit que les activités moyennes de différentes populations de neurones sans pouvoir étudier leurs corrélations, et beaucoup d'efforts sont mis dans la recherche actuelle afin d'en inclure. Par ailleurs, le modèle de Wilson–Cowan est très semblable à différents modèles utilisés dans d'autres disciplines, comme les modèles épidémiologiques compartimentaux. Malgré les similarités qui existent entre ces modèles, leurs liens sont peu étudiés dans la littérature. Dans ce mémoire, on voulait ainsi proposer une nouvelle construction plus générale de ces modèles, dont chacun serait un cas particulier. On espérait ainsi mieux comprendre leurs fondements mathématiques, et parvenir à proposer une généralisation du modèle de Wilson–Cowan qui permettrait de décrire des corrélations entre les états de différents neurones.

Dans le but de modéliser des phénomènes biologiques où les unités peuvent prendre trois états qui évoluent de façon cyclique, on a ainsi proposé au chapitre 2 la construction d'un processus stochastique qui attribue des états aux nœuds d'un graphe et qui décrit leur évolution. On a ensuite vu comment ce modèle microscopique peut mener à une description macroscopique de la dynamique sur le graphe. En effet, on a supposé que les nœuds étaient répartis en un certain nombre de populations, et on a défini des variables aléatoires qui représentent les fractions des nœuds de chaque population qui se trouvent dans chacun des trois états possibles. On a alors décrit une méthode, basée sur un problème de fermeture de moments, qui permet d'obtenir un modèle macroscopique sous la forme d'un système d'équations différentielles ordinaires. Un tel système peut alors inclure les seconds moments des variables aléatoires qui correspondent à l'état macroscopique comme il peut décrire seulement l'évolution de leurs valeurs moyennes.

On a ensuite présenté deux applications de cette modélisation. Au chapitre 3, on a vu une application en épidémiologie, qui mène à une nouvelle construction du modèle SIR. Le problème de fermeture de moments pour ce modèle a déjà été largement étudié dans la littérature. On a donc présenté quelques solutions qui ont déjà été proposées, et qui incluent ou non les seconds moments des variables aléatoires d'intérêt.

Ensuite, on a vu au chapitre 4 une application en neurosciences, qui permet de modéliser l'activité d'un réseau de neurones biologiques. Dans ce contexte, on a obtenu un problème

de fermeture de moments dont les solutions décrivent la dynamique du réseau de neurones d'un point de vue macroscopique. À la connaissance de l'auteur, ce problème n'avait jamais été étudié précédemment. Cependant, on a vu que sa solution de champ moyen correspond à une légère généralisation du modèle de Wilson–Cowan. En effet, un système équivalent à celui de Wilson–Cowan peut être obtenu à partir du système de champ moyen en y fixant les fractions réfractaires des populations de neurones à leurs états d'équilibre, qui sont donnés en fonction des fractions actives des populations. Une comparaison entre ces deux modèles a alors montré que le modèle équivalent à celui de Wilson–Cowan ne peut pas représenter tous les comportements prévus par le modèle de champ moyen, qui inclut l'état réfractaire explicitement.

Après avoir étudié ce modèle de champ moyen, on a étudié d'autres solutions au problème de fermeture de moments pour ce modèle de dynamique neuronale afin d'inclure les moments d'ordre 2. Deux nouvelles solutions ont alors été proposées. Malheureusement, l'étude de la première de ces deux solutions n'a pas montré que ce modèle permet d'améliorer les prédictions du modèle de champ moyen. Au contraire, on a vu des cas où ce modèle prévoit des solutions invraisemblables, et d'autres cas où ses prédictions ne correspondent pas au comportement du processus stochastique sous-jacent alors que celles du modèle de champ moyen en sont une bonne approximation. Cependant, la seconde solution qu'on a proposée a plus de potentiel. En effet, on a vu que ce modèle ne souffre pas des problèmes de la première solution, et on a vu un exemple où il permet d'améliorer les prédictions du modèle de champ moyen : dans ce cas, la solution de ce modèle est une meilleure approximation des trajectoires du processus stochastique sous-jacent que la solution du modèle de champ moyen. Ce nouveau modèle n'a toutefois pu être étudié que par quelques exemples. Il serait donc intéressant d'en faire une étude théorique plus poussée lors de travaux futurs.

L'étude de ces applications a également mis en évidence plusieurs défis qui surviennent lorsqu'on tente de trouver un système autonome d'équations différentielles à partir du problème de fermeture de moments. En effet, les nouvelles constructions de modèles classiques ont permis de mieux en cerner les fondements mathématiques, et ainsi de mieux en comprendre les limites.

En particulier, on a vu que la construction des modèles macroscopiques dépend de l'existence des taux de transition moyens de la forme (2.59). Or, dans les applications qu'on a étudiées, l'existence de tels taux de transition moyens n'est garantie que lorsque les paramètres qui caractérisent les nœuds du graphe sont constants dans une même population. Si ce n'est pas le cas, les taux de transition moyens ne sont qu'approximatifs. Toutefois, on a vu que si les paramètres sont modélisés comme des variables aléatoires indépendantes et identiquement distribuées, ces approximations sont justifiées, à tout le moins lorsqu'il n'y a pas de corrélation entre les paramètres et les états des nœuds. Dans le cas contraire, les approximations sont moins bonnes. Par exemple, dans le contexte du modèle de dynamique

neuronale étudié au chapitre 4, il est logique d'imaginer qu'il peut y avoir une corrélation entre l'état d'un neurone et son seuil d'activation. Ceci pourrait expliquer pourquoi il arrive que les prédictions des modèles macroscopiques ne concordent pas avec le comportement du processus stochastique sous-jacent, comme c'est le cas dans l'exemple discuté à la section 4.3.2. Ainsi, il serait intéressant de voir si un processus stochastique peut servir à construire un modèle macroscopique à partir d'une méthode différente de celle proposée à la section 2.4.

Une autre faiblesse majeure des modélisations qu'on a étudiées dans ce mémoire est la difficulté à caractériser dans quelle mesure une solution à un problème de fermeture de moments est approximative. Dans tous les cas étudiés ici, la seule méthode qu'on a pu utiliser afin de vérifier si un système macroscopique est cohérent avec le processus stochastique sous-jacent est de comparer les solutions du système à des trajectoires du processus stochastique simulées numériquement.

Finalement, on remarque que, malgré le fait que la dynamique markovienne ternaire cyclique construite au chapitre 2 puisse être appliquée dans plusieurs contextes afin de représenter différents phénomènes biologiques, la forme du processus stochastique en limite forcément les champs d'application. Or, il serait possible de construire un processus stochastique un peu plus général, notamment en permettant davantage d'états possibles. Les méthodes qu'on a utilisées dans ce mémoire pourraient alors être appliquées pour retrouver, par exemple, les autres modèles épidémiologiques compartimentaux, ou pour modéliser bien d'autres phénomènes biologiques.

Bibliographie

- [1] L. J. S. ALLEN. « An Introduction to Stochastic Epidemic Models ». In : *Mathematical Epidemiology*. Sous la dir. de F. BRAUER, P. van den DRIESSCHE et J. Wu. Lecture Notes in Mathematics. Springer, 2008, p. 81-130. ISBN : 978-3-540-78911-6. doi : 10.1007/978-3-540-78911-6_3.
- [2] M. BENAYOUN *et al.* « Avalanches in a Stochastic Model of Spiking Neurons ». In : *PLOS Computational Biology* 6.7 (2010), e1000846. ISSN : 1553-7358. doi : 10.1371/journal.pcbi.1000846.
- [3] R. L. BEURLE. « Properties of a Mass of Cells Capable of Regenerating Pulses ». In : *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 240.669 (1956), p. 55-94. doi : 10.2307/92540.
- [4] R. BHATIA et C. DAVIS. « A Better Bound on the Variance ». In : *The American Mathematical Monthly* 107.4 (2000), p. 353-357. doi : 10.1080/00029890.2000.12005203.
- [5] B. BOLLOBÁS. *Random Graphs*. 2^e éd. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2001. ISBN : 978-0-521-80920-7. doi : 10.1017/CBO9780511814068.
- [6] J.-M. BONY. « Principe du maximum, inégalité de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés ». In : *Annales de l'institut Fourier* 19.1 (1969), p. 277-304. doi : 10.5802/aif.319.
- [7] F. BRAUER. « Compartmental Models in Epidemiology ». In : *Mathematical Epidemiology*. Sous la dir. de F. BRAUER, P. van den DRIESSCHE et J. Wu. Lecture Notes in Mathematics. Springer, 2008, p. 19-79. ISBN : 978-3-540-78911-6. doi : 10.1007/978-3-540-78911-6_2.
- [8] P. C. BRESSLOFF. « Metastable states and quasicycles in a stochastic Wilson-Cowan model of neuronal population dynamics ». In : *Physical Review E* 82.5 (2010), p. 051903. ISSN : 15393755. doi : 10.1103/PhysRevE.82.051903.
- [9] P. C. BRESSLOFF. « Stochastic Neural Field Theory and the System-Size Expansion ». In : *SIAM Journal on Applied Mathematics* 70.5 (2009), p. 1488-1521. ISSN : 0036-1399. doi : 10.1137/090756971.
- [10] P. C. BRESSLOFF *et al.* « Stochastic Network Models in Neuroscience : A Festschrift for Jack Cowan. Introduction to the Special Issue ». In : *Journal of Mathematical Neuroscience* 6.1 (2016), p. 4. ISSN : 21908567. doi : 10.1186/s13408-016-0036-y.

- [11] H. BREZIS. « On a characterization of flow-invariant sets ». In : *Communications on Pure and Applied Mathematics* 23.2 (1970), p. 261-263. doi : 10.1002/cpa.3160230211.
- [12] M. A. BUICE et J. D. COWAN. « Field-theoretic approach to fluctuation effects in neural networks ». In : *Physical Review E* 75.5 (2007), p. 051919. doi : 10.1103/PhysRevE.75.051919.
- [13] M. A. BUICE et J. D. COWAN. « Statistical mechanics of the neocortex ». In : *Progress in Biophysics and Molecular Biology* 99.2 (2009), p. 53-86. issn : 0079-6107. doi : 10.1016/j.pbiomolbio.2009.07.003.
- [14] M. A. BUICE, J. D. COWAN et C. C. CHOW. « Systematic fluctuation expansion for neural network activity equations ». In : *Neural Computation* 22.2 (2010), p. 377-426. issn : 08997667. doi : 10.1162/neco.2009.02-09-960.
- [15] C.-R. CAI et Z.-X. WU. « Analytical treatment for cyclic three-state dynamics on static networks ». In : *Physical Review E* 101.1 (2020), p. 012305. doi : 10.1103/PhysRevE.101.012305.
- [16] E. CATOR et P. VAN MIEGHEM. « Second-order mean-field susceptible-infected-susceptible epidemic threshold ». In : *Physical Review E* 85.5 (2012), p. 056111. doi : 10.1103/PhysRevE.85.056111.
- [17] C. C. CHOW et Y. KARIMIPANAH. « Before and beyond the Wilson–Cowan equations ». In : *Journal of Neurophysiology* 123.5 (2020), p. 1645-1656. doi : 10.1152/jn.00404.2019.
- [18] E. ÇINLAR. *Probability and stochastics*. T. 261. Graduate Texts in Mathematics. Springer-Verlag, 2011. isbn : 978-0-387-87858-4.
- [19] J. D. COWAN. « A Personal Account of the Development of the Field Theory of Large-Scale Brain Activity from 1945 Onward ». In : *Neural Fields : Theory and Applications*. Sous la dir. de S. COOMBES *et al.* Springer, 2014, p. 47-96. isbn : 978-3-642-54593-1. doi : 10.1007/978-3-642-54593-1_2.
- [20] J. D. COWAN. « Stochastic Neurodynamics ». In : *Advances in Neural Information Processing Systems*. T. 3. Morgan-Kaufmann, 1990, p. 62-69.
- [21] J. D. COWAN, J. NEUMAN et W. van DRONGELEN. « Wilson–Cowan Equations for Neocortical Dynamics ». In : *The Journal of Mathematical Neuroscience* 6.1 (2016), p. 1. doi : 10.1186/s13408-015-0034-5.
- [22] R. CURTU et B. ERMENTROUT. « Oscillations in a refractory neural net ». In : *Journal of Mathematical Biology* 43.1 (2001), p. 81-100. issn : 1432-1416. doi : 10.1007/s002850100089.
- [23] R. W. R. DARLING et J. R. NORRIS. « Differential equation approximations for Markov chains ». In : *Probability Surveys* 5 (2008), p. 37-79. issn : 1549-5787. doi : 10.1214/07-PS121.

- [24] A. DESTEXHE et T. J. SEJNOWSKI. « The Wilson-Cowan model, 36 years later ». In : *Biological Cybernetics* 101 (2009), p. 1-2. doi : 10.1007/s00422-009-0328-3.
- [25] J. L. DOOB. *Stochastic Processes*. Wiley Classics Library. John Wiley & Sons, 1990. 664 p. ISBN : 978-0-471-52369-7.
- [26] D. J. D. EARN. « A Light Introduction to Modelling Recurrent Epidemics ». In : *Mathematical Epidemiology*. Sous la dir. de F. BRAUER, P. van den DRIESSCHE et J. WU. Lecture Notes in Mathematics. Springer, 2008, p. 3-17. ISBN : 978-3-540-78911-6. doi : 10.1007/978-3-540-78911-6_1.
- [27] D. J. D. EARN. « Mathematical epidemiology of infectious diseases ». In : *Mathematical Biology*. IAS/Park City Mathematics Series. American Mathematical Society, 2009, p. 151-186.
- [28] S. EL BOUSTANI et A. DESTEXHE. « A master equation formalism for macroscopic modeling of asynchronous irregular activity states ». In : *Neural Computation* 21.1 (2009), p. 46-100. doi : 10.1162/neco.2009.02-08-710.
- [29] G. B. ERMENTROUT et J. D. COWAN. « A mathematical theory of visual hallucination patterns ». In : *Biological Cybernetics* 34.3 (1979), p. 137-150. ISSN : 1432-0770. doi : 10.1007/BF00336965.
- [30] G. B. ERMENTROUT et J. D. COWAN. « Large Scale Spatially Organized Activity in Neural Nets ». In : *SIAM Journal on Applied Mathematics* 38.1 (1980), p. 1-21. ISSN : 0036-1399. doi : 10.1137/0138001.
- [31] G. B. ERMENTROUT et J. D. COWAN. « Temporal oscillations in neuronal nets ». In : *Journal of Mathematical Biology* 7.3 (1979), p. 265-280. ISSN : 1432-1416. doi : 10.1007/BF00275728.
- [32] G. B. ERMENTROUT et D. H. TERMAN. *Mathematical Foundations of Neuroscience*. T. 35. Interdisciplinary Applied Mathematics. Springer, 2010. ISBN : 978-0-387-87707-5. doi : 10.1007/978-0-387-87708-2.
- [33] O. D. FAUGERAS, J. D. TOBOUL et B. CESSAC. « A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs ». In : *Frontiers in Computational Neuroscience* 3 (2009). ISSN : 1662-5188. doi : 10.3389/neuro.10.001.2009.
- [34] L. FRACHEBOURG, P. L. KRAPIVSKY et E. BEN-NAIM. « Spatial organization in cyclic Lotka-Volterra systems ». In : *Physical Review E* 54.6 (1996), p. 6186-6200. doi : 10.1103/PhysRevE.54.6186.
- [35] K.-i. FUNAHASHI et Y. NAKAMURA. « Approximation of dynamical systems by continuous time recurrent neural networks ». In : *Neural Networks* 6.6 (1993), p. 801-806. ISSN : 0893-6080. doi : 10.1016/S0893-6080(05)80125-X.

- [36] S. A. GERSHGORIN. « Über die Abgrenzung der Eigenwerte einer Matrix ». In : *Izv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk* 6 (1931), p. 749-754.
- [37] D. T. GILLESPIE. « A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions ». In : *Journal of Computational Physics* 22 (1976), p. 403-434. doi : [10.1016/0021-9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3).
- [38] I. GOYCHUK et A. GOYCHUK. « Stochastic Wilson–Cowan models of neuronal network dynamics with memory and delay ». In : *New Journal of Physics* 17.4 (2015), p. 045029. doi : [10.1088/1367-2630/17/4/045029](https://doi.org/10.1088/1367-2630/17/4/045029).
- [39] J. S. GRIFFITH. « A field theory of neural nets : I : Derivation of field equations ». In : *The Bulletin of Mathematical Biophysics* 25.1 (1963), p. 111-120. doi : [10.1007/BF02477774](https://doi.org/10.1007/BF02477774).
- [40] J. S. GRIFFITH. « A field theory of neural nets : II. Properties of the field equations ». In : *The Bulletin of Mathematical Biophysics* 27.2 (1965), p. 187-195. doi : [10.1007/BF02498774](https://doi.org/10.1007/BF02498774).
- [41] D. M. GROBMAN. « Homeomorphisms of systems of differential equations ». In : *Doklady Akademii Nauk SSSR* 128.5 (1959), p. 880-881.
- [42] S. GROSSBERG. « Nonlinear neural networks : Principles, mechanisms, and architectures ». In : *Neural Networks* 1.1 (1988), p. 17-61. issn : 0893-6080. doi : [10.1016/0893-6080\(88\)90021-4](https://doi.org/10.1016/0893-6080(88)90021-4).
- [43] J. GUCKENHEIMER et P. J. HOLMES. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Applied Mathematical Sciences. Springer-Verlag, 1983. isbn : 978-0-387-90819-9. doi : [10.1007/978-1-4612-1140-2](https://doi.org/10.1007/978-1-4612-1140-2).
- [44] P. R. HALMOS. *Measure Theory*. T. 18. Graduate Texts in Mathematics. Springer-Verlag, 1950. 304 p. isbn : 978-0-387-90088-9. doi : [10.1007/978-1-4684-9440-2](https://doi.org/10.1007/978-1-4684-9440-2).
- [45] P. HARTMAN. « A Lemma in the Theory of Structural Stability of Differential Equations ». In : *Proceedings of the American Mathematical Society* 11.4 (1960), p. 610-620. issn : 0002-9939. doi : [10.2307/2034720](https://doi.org/10.2307/2034720).
- [46] H. W. HETHCOTE. « The Mathematics of Infectious Diseases ». In : *SIAM Review* 42.4 (2000), p. 599-653. issn : 0036-1445. doi : [10.1137/S0036144500371907](https://doi.org/10.1137/S0036144500371907).
- [47] M. W. HIRSCH. « Convergent activation dynamics in continuous time networks ». In : *Neural Networks* 2.5 (1989), p. 331-349. issn : 0893-6080. doi : [10.1016/0893-6080\(89\)90018-X](https://doi.org/10.1016/0893-6080(89)90018-X).
- [48] J. J. HOPFIELD. « Neurons with graded response have collective computational properties like those of two-state neurons ». In : *Proceedings of the National Academy of Sciences* 81.10 (1984), p. 3088-3092. issn : 0027-8424, 1091-6490. doi : [10.1073/pnas.81.10.3088](https://doi.org/10.1073/pnas.81.10.3088).

-
- [49] F. C. HOPPENSTEADT et E. M. IZHIKEVICH. *Weakly Connected Neural Networks*. Applied Mathematical Sciences. Springer-Verlag, 1997. ISBN : 978-0-387-94948-2. doi : 10.1007/978-1-4612-1828-9.
 - [50] L. HÖRMANDER. *The Analysis of Linear Partial Differential Operators I : Distribution Theory and Fourier Analysis*. 2^e éd. Classics in Mathematics, The Analysis of Linear Partial Differential Operators. Springer-Verlag, 2003. ISBN : 978-3-540-00662-6. doi : 10.1007/978-3-642-61497-2.
 - [51] R. A. HORN et C. R. JOHNSON. *Matrix Analysis*. 2^e éd. Cambridge University Press, 2012. 662 p. ISBN : 978-0-521-83940-2.
 - [52] T. HOUSE *et al.* « A Motif-Based Approach to Network Epidemics ». In : *Bulletin of Mathematical Biology* 71.7 (2009), p. 1693-1706. ISSN : 1522-9602. doi : 10.1007/s11538-009-9420-z.
 - [53] J. D. HUNTER. « Matplotlib : A 2D Graphics Environment ». In : *Computing in Science Engineering* 9.3 (2007), p. 90-95. ISSN : 1558-366X. doi : 10.1109/MCSE.2007.55.
 - [54] J. JOO et J. L. LEBOWITZ. « Pair approximation of the stochastic susceptible-infected-recovered-susceptible epidemic model on the hypercubic lattice ». In : *Physical Review E* 70.3 (2004), p. 036114. doi : 10.1103/PhysRevE.70.036114.
 - [55] M. J. KEELING. « The effects of local spatial structure on epidemiological invasions ». In : *Proceedings of the Royal Society of London. Series B : Biological Sciences* 266.1421 (1999), p. 859-867. doi : 10.1098/rspb.1999.0716.
 - [56] M. J. KEELING, D. A. RAND et A. J. MORRIS. « Correlation models for childhood epidemics ». In : *Proceedings of the Royal Society of London. Series B : Biological Sciences* 264.1385 (1997), p. 1149-1156. doi : 10.1098/rspb.1997.0159.
 - [57] W. O. KERMACK et A. G. MCKENDRICK. « A contribution to the mathematical theory of epidemics ». In : *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 115.772 (1927), p. 700-721. doi : 10.1098/rspa.1927.0118.
 - [58] W. O. KERMACK et A. G. MCKENDRICK. « Contributions to the mathematical theory of epidemics. II. —The problem of endemicity ». In : *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 138.834 (1932), p. 55-83. doi : 10.1098/rspa.1932.0171.
 - [59] W. O. KERMACK et A. G. MCKENDRICK. « Contributions to the mathematical theory of epidemics. III.—Further studies of the problem of endemicity ». In : *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 141.843 (1933), p. 94-122. doi : 10.1098/rspa.1933.0106.

-
- [60] I. Z. KISS, J. C. MILLER et P. L. SIMON. *Mathematics of Epidemics on Networks*. T. 46. Interdisciplinary Applied Mathematics September 2005. Springer International Publishing, 2017. 423 p. ISBN : 978-3-319-50804-7. DOI : [10.1007/978-3-319-50806-1](https://doi.org/10.1007/978-3-319-50806-1).
 - [61] I. Z. KISS *et al.* « Exact deterministic representation of Markovian SIR epidemics on networks with and without loops ». In : *Journal of Mathematical Biology* 70.3 (2015), p. 437-464. ISSN : 1432-1416. DOI : [10.1007/s00285-014-0772-0](https://doi.org/10.1007/s00285-014-0772-0).
 - [62] A. KLENKE. *Probability Theory : A Comprehensive Course*. 3^e éd. Universitext. Springer International Publishing, 2020. ISBN : 978-3-030-56401-8. DOI : [10.1007/978-3-030-56402-5](https://doi.org/10.1007/978-3-030-56402-5).
 - [63] A. N. KOLMOGOROV. *Foundations of the Theory of Probability*. Trad. par N. MORRISON. Martino Fine Books, 2013. 96 p. ISBN : 978-1-61427-514-5.
 - [64] A. N. KOLMOGOROV. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Ergebnisse der Mathematik und Ihrer Grenzgebiete. 1. Folge. Springer-Verlag, 1933. ISBN : 978-3-642-49596-0. DOI : [10.1007/978-3-642-49888-6](https://doi.org/10.1007/978-3-642-49888-6).
 - [65] C. KUEHN. « Moment Closure—A Brief Review ». In : *Control of Self-Organizing Non-linear Systems*. Sous la dir. d'E. SCHÖLL, S. H. L. KLAPP et P. HÖVEL. Understanding Complex Systems. Springer International Publishing, 2016, p. 253-271. ISBN : 978-3-319-28028-8. DOI : [10.1007/978-3-319-28028-8_13](https://doi.org/10.1007/978-3-319-28028-8_13).
 - [66] T. G. KURTZ. « Solutions of ordinary differential equations as limits of pure jump markov processes ». In : *Journal of Applied Probability* 7.1 (1970), p. 49-58. ISSN : 0021-9002, 1475-6072. DOI : [10.2307/3212147](https://doi.org/10.2307/3212147).
 - [67] S. A. LEVIN et R. DURRETT. « From individuals to epidemics ». In : *Philosophical Transactions of the Royal Society of London. Series B : Biological Sciences* 351.1347 (1996), p. 1615-1621. DOI : [10.1098/rstb.1996.0145](https://doi.org/10.1098/rstb.1996.0145).
 - [68] W. A. LITTLE. « The existence of persistent states in the brain ». In : *Mathematical Biosciences* 19.1 (1974), p. 101-120. ISSN : 0025-5564. DOI : [10.1016/0025-5564\(74\)90031-5](https://doi.org/10.1016/0025-5564(74)90031-5).
 - [69] A. J. LOTKA. « Analytical Note on Certain Rhythmic Relations in Organic Systems ». In : *Proceedings of the National Academy of Sciences of the United States of America* 6.7 (1920), p. 410-415. ISSN : 0027-8424.
 - [70] M. MARTCHEVA. *An Introduction to Mathematical Epidemiology*. Texts in Applied Mathematics. Springer US, 2015. ISBN : 978-1-4899-7611-6. DOI : [10.1007/978-1-4899-7612-3](https://doi.org/10.1007/978-1-4899-7612-3).
 - [71] H. MATSUDA *et al.* « Statistical Mechanics of Population : The Lattice Lotka-Volterra Model ». In : *Progress of Theoretical Physics* 88.6 (1992), p. 1035-1049. ISSN : 0033-068X. DOI : [10.1143/ptp/88.6.1035](https://doi.org/10.1143/ptp/88.6.1035).

- [72] W. S. McCULLOCH et W. PITTS. « A logical calculus of the ideas immanent in nervous activity ». In : *The Bulletin of Mathematical Biophysics* 5.4 (1943), p. 115-133. ISSN : 00074985. DOI : 10.1007/BF02478259.
- [73] M. NAGUMO. « Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen ». In : *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series* 24 (1942), p. 551-559. DOI : 10.11429/ppmsj1919.24.0_551.
- [74] J. R. NORRIS. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997. ISBN : 978-0-521-63396-3. DOI : 10.1017/CBO9780511810633.
- [75] C. NOWZARI, V. M. PRECIADO et G. J. PAPPAS. « Analysis and Control of Epidemics : A Survey of Spreading Processes on Complex Networks ». In : *IEEE Control Systems Magazine* 36.1 (2016), p. 26-46. ISSN : 1941-000X. DOI : 10.1109/MCS.2015.2495000.
- [76] A. OKUBO et S. A. LEVIN. *Diffusion and Ecological Problems : Modern Perspectives*. 2^e éd. Interdisciplinary Applied Mathematics. Springer-Verlag, 2001. ISBN : 978-0-387-98676-0. DOI : 10.1007/978-1-4757-4978-6.
- [77] V. PAINCHAUD. *PopNet*. Version 1.0.0. 2021. URL : <https://github.com/vincentpainchaud/PopNet>.
- [78] R. PASTOR-SATORRAS *et al.* « Epidemic processes in complex networks ». In : *Reviews of Modern Physics* 87.3 (2015), p. 925-979. DOI : 10.1103/RevModPhys.87.925.
- [79] L. PELLIS, T. HOUSE et M. J. KEELING. « Exact and approximate moment closures for non-Markovian network epidemics ». In : *Journal of Theoretical Biology* 382 (2015), p. 160-177. ISSN : 0022-5193. DOI : 10.1016/j.jtbi.2015.04.039.
- [80] L. PERKO. *Differential Equations and Dynamical Systems*. 3^e éd. Texts in Applied Mathematics. Springer-Verlag, 2001. ISBN : 978-0-387-95116-4. DOI : 10.1007/978-1-4613-0003-8.
- [81] T. POPOVICIU. « Sur les équations algébriques ayant toutes leurs racines réelles ». In : *Mathematica* 9 (1935), p. 129-145.
- [82] M. A. PORTER et J. P. GLEESON. *Dynamical Systems on Networks : A Tutorial*. Frontiers in Applied Dynamical Systems : Reviews and Tutorials. Springer International Publishing, 2016. ISBN : 978-3-319-26640-4. DOI : 10.1007/978-3-319-26641-1.
- [83] D. A. RAND. « Correlation Equations and Pair Approximations for Spatial Ecologies ». In : *Advanced Ecological Theory*. John Wiley & Sons, Ltd, 1999, p. 100-142. ISBN : 978-1-4443-1150-1.
- [84] T. REICHENBACH, M. MOBILIA et E. FREY. « Coexistence versus extinction in the stochastic cyclic Lotka-Volterra model ». In : *Physical Review E* 74.5 (2006), p. 051907. DOI : 10.1103/PhysRevE.74.051907.

- [85] T. ROGERS. « Maximum-entropy moment-closure for stochastic systems on networks ». In : *Journal of Statistical Mechanics : Theory and Experiment* 2011.5 (2011), P05007. issn : 1742-5468. doi : 10.1088/1742-5468/2011/05/P05007.
- [86] W. RUDIN. *Real and Complex Analysis*. 3^e éd. Series in Higher Mathematics. McGraw-Hill, 1987. isbn : 978-0-07-054234-1.
- [87] M. E. RULE *et al.* « Neural field models for latent state inference : Application to large-scale neuronal recordings ». In : *PLOS Computational Biology* 15.11 (2019), e1007442. issn : 1553-7358. doi : 10.1371/journal.pcbi.1007442.
- [88] K. SATŌ, H. MATSUDA et A. SASAKI. « Pathogen invasion and host extinction in lattice structured populations ». In : *Journal of Mathematical Biology* 32.3 (1994), p. 251-268. issn : 1432-1416. doi : 10.1007/BF00163881.
- [89] D. G. SCHAEFFER et J. W. CAIN. *Ordinary Differential Equations : Basics and Beyond*. T. 65. Texts in Applied Mathematics. Springer-Verlag, 2016. 542 p. isbn : 978-1-4939-6387-4.
- [90] K. J. SHARKEY. « Deterministic epidemic models on contact networks : Correlations and unbiological terms ». In : *Theoretical Population Biology* 79.4 (2011), p. 115-129. issn : 0040-5809. doi : 10.1016/j.tpb.2011.01.004.
- [91] K. J. SHARKEY. « Deterministic epidemiological models at the individual level ». In : *Journal of Mathematical Biology* 57.3 (2008), p. 311-331. issn : 1432-1416. doi : 10.1007/s00285-008-0161-7.
- [92] A. SHIMBEL et A. RAPOPORT. « A statistical approach to the theory of the central nervous system ». In : *The Bulletin of Mathematical Biophysics* 10.1 (1948), p. 41-55. doi : 10.1007/BF02478329.
- [93] J. STARK, P. IANNELLI et S. BAIGENT. « A nonlinear dynamics perspective of moment closure for stochastic processes ». In : *Nonlinear Analysis : Theory, Methods & Applications*. Proceedings of the Third World Congres of Nonlinear Analysts 47.2 (2001), p. 753-764. issn : 0362-546X. doi : 10.1016/S0362-546X(01)00220-6.
- [94] P. VAN MIEGHEM. « The N-intertwined SIS epidemic network model ». In : *Computing* 93.2 (2011), p. 147-169. issn : 1436-5057. doi : 10.1007/s00607-011-0155-y.
- [95] P. VAN MIEGHEM, J. OMIC et R. KOORIJ. « Virus Spread in Networks ». In : *IEEE/ACM Transactions on Networking* 17.1 (2009), p. 1-14. doi : 10.1109/TNET.2008.925623.
- [96] P. VIRTANEN *et al.* « SciPy 1.0 : fundamental algorithms for scientific computing in Python ». In : *Nature Methods* 17.3 (2020), p. 261-272. issn : 1548-7105. doi : 10.1038/s41592-019-0686-2.
- [97] T. P. VOGELS, K. RAJAN et L. ABBOTT. « Neural network dynamics ». In : *Annual Review of Neuroscience* 28.1 (2005), p. 357-376. issn : 0147-006X. doi : 10.1146/annurev.neuro.28.061604.135637.

-
- [98] V. VOLTERRA. « Variazioni e fluttuazioni del numero d'individui in specie animali conviventi ». In : *Accademia dei Lincei* 2 (1926), p. 31-113.
 - [99] C. WEISTUCH, L. R. MUJICA-PARODI et K. DILL. « The Refractory Period Matters : Unifying Mechanisms of Macroscopic Brain Waves ». In : *Neural Computation* 33.5 (2021), p. 1145-1163. ISSN : 0899-7667. DOI : [10.1162/neco_a_01371](https://doi.org/10.1162/neco_a_01371).
 - [100] S. WIGGINS. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. 2^e éd. T. 2. Texts in Applied Mathematics. Springer-Verlag, 2003. ISBN : 978-0-387-00177-7. DOI : [10.1007/b97481](https://doi.org/10.1007/b97481).
 - [101] H. R. WILSON et J. D. COWAN. « Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons ». In : *Biophysical Journal* 12.1 (1972), p. 1-24. DOI : [10.1016/S0006-3495\(72\)86068-5](https://doi.org/10.1016/S0006-3495(72)86068-5).
 - [102] H. WOLKOWICZ et G. P. STYAN. « Bounds for eigenvalues using traces ». In : *Linear Algebra and Its Applications* 29 (1980), p. 471-506. DOI : [10.1016/0024-3795\(80\)90258-X](https://doi.org/10.1016/0024-3795(80)90258-X).
 - [103] H. WOLKOWICZ et G. P. STYAN. « More bounds for eigenvalues using traces ». In : *Linear Algebra and Its Applications* 31 (1980), p. 1-17. DOI : [10.1016/0024-3795\(80\)90201-3](https://doi.org/10.1016/0024-3795(80)90201-3).
 - [104] M. ZAREPOUR *et al.* « Universal and nonuniversal neural dynamics on small world connectomes : A finite-size scaling analysis ». In : *Physical Review E* 100.5 (2019), p. 52138. ISSN : 24700053. DOI : [10.1103/PhysRevE.100.052138](https://doi.org/10.1103/PhysRevE.100.052138).
 - [105] J. ZHANG *et al.* « Dynamics analysis of SIR epidemic model with correlation coefficients and clustering coefficient in networks ». In : *Journal of Theoretical Biology* 449 (2018), p. 1-13. ISSN : 0022-5193. DOI : [10.1016/j.jtbi.2018.04.007](https://doi.org/10.1016/j.jtbi.2018.04.007).

Annexe A

Dérivées des seconds moments

Ce chapitre présente les démarches nécessaires afin d'obtenir les systèmes (2.61) et (2.62), qui décrivent comment évoluent les covariances de fractions de populations d'un graphe aléatoire sur lequel une dynamique MTC est définie. Pour dériver ces systèmes, la démarche est très semblable à celle présentée au chapitre 2 pour obtenir le système (2.60). Dans les preuves présentées ici, il arrivera fréquemment que plusieurs calculs soient très similaires entre eux. Afin d'éviter d'alourdir inutilement ce chapitre, on présente souvent les détails pour un seul cas.

On considère G un graphe aléatoire d'ordre N sur un espace de probabilité $(\Theta, \mathcal{G}, \mu)$. On suppose qu'une dynamique MTC est définie pour G sur l'espace $(\Theta \times \Omega, \mathcal{G} \otimes \mathcal{F}, \mathbb{P})$, et qu'elle est décrite par les taux a_j , b_j et c_j , au sens de la définition 2.10. On note $\{X_t\}_{t \geq 0}$ le processus de projection de coordonnées sur (Ω, \mathcal{F}) , et on rappelle que l'on interprète une composante X_t^j de X_t comme l'état du nœud j au temps t .

Pour chaque paire de nœuds j, k et chaque $\vartheta \in \Theta$, soient $p_{jk}^\vartheta, r_{jk}^\vartheta, \rho_{jk}^\vartheta: [0, \infty) \rightarrow [0, 1]$ données par

$$p_{jk}^\vartheta(t) := \mathbb{P}[X_t^j = X_t^k = 1 \mid \vartheta], \quad r_{jk}^\vartheta(t) := \mathbb{P}[X_t^j = X_t^k = i \mid \vartheta] \quad (\text{A.1a})$$

et

$$\rho_{jk}^\vartheta(t) := \mathbb{P}[X_t^j = 1, X_t^k = i \mid \vartheta]. \quad (\text{A.1b})$$

On introduit également les fonctions $p_{jk}, r_{jk}, \rho_{jk}: [0, \infty) \rightarrow [0, 1]$ données par

$$p_{jk}(t) := \mathbb{P}[X_t^j = X_t^k = 1], \quad r_{jk}(t) := \mathbb{P}[X_t^j = X_t^k = i], \quad (\text{A.2a})$$

et

$$\rho_{jk}(t) := \mathbb{P}[X_t^j = 1, X_t^k = i]. \quad (\text{A.2b})$$

On introduit d'abord le résultat suivant pour manipuler plus facilement ces probabilités.

Lemme A.1. Soit G un graphe aléatoire muni d'une dynamique MTC. Pour chaque paire de nœuds j, k de G , soient $p_{jk}^\vartheta, r_{jk}^\vartheta, \rho_{jk}^\vartheta: [0, \infty) \rightarrow [0, 1]$ définies par (A.1). Alors pour tout $t \geq 0$ et pour tout $\vartheta \in \Theta$, $p_{jk}^\vartheta(t) = \mathbb{E}[\operatorname{Re}(X_t^j) \operatorname{Re}(X_t^k) \mid \vartheta]$, $r_{jk}^\vartheta(t) = \mathbb{E}[\operatorname{Im}(X_t^j) \operatorname{Im}(X_t^k) \mid \vartheta]$ et $\rho_{jk}^\vartheta(t) = \mathbb{E}[\operatorname{Re}(X_t^j) \operatorname{Im}(X_t^k) \mid \vartheta]$. ◆

Démonstration. Ces résultats suivent tout simplement du fait que chaque X_t^j prend ses valeurs dans l'ensemble $\{0, 1, i\}$. En effet, soit $t \geq 0$, une paire de nœuds j, k de G , et $\vartheta \in \Theta$. Alors

$$\rho_{jk}^\vartheta(t) = \mathbb{P}[X_t^j = 1, X_t^k = i | \vartheta] = \text{Re}(1) \text{Im}(i) \mathbb{P}[X_t^j = 1, X_t^k = i] \quad (\text{A.3})$$

$$= \sum_{x,y \in \{0,1,i\}} \text{Re}(x) \text{Im}(y) \mathbb{P}[X_t^j = x, X_t^k = y | \vartheta] = \mathbb{E}[\text{Re}(X_t^j) \text{Im}(X_t^k) | \vartheta]. \quad (\text{A.4})$$

Les résultats pour p_{jk}^ϑ et r_{jk}^ϑ sont obtenus de la même façon, en faisant apparaître à la seconde égalité deux parties réelles ou bien deux parties imaginaires selon le cas. ■

L'analogue de ce résultat sans conditionnement sur la projection sur l'espace des paramètres est aussi vrai.

Lemme A.2. Soit G un graphe aléatoire muni d'une dynamique MTC. Pour chaque paire de nœuds j, k de G , soient $p_{jk}, r_{jk}, \rho_{jk}: [0, \infty) \rightarrow [0, 1]$ définies par (A.2). Alors pour tout $t \geq 0$, $p_{jk}(t) = \mathbb{E}[\text{Re}(X_t^j) \text{Re}(X_t^k)]$, $r_{jk}(t) = \mathbb{E}[\text{Im}(X_t^j) \text{Im}(X_t^k)]$ et $\rho_{jk}(t) = \mathbb{E}[\text{Re}(X_t^j) \text{Im}(X_t^k)]$. ◆

Démonstration. Soient $t \geq 0$ et j, k une paire de nœuds de G . On a vu au lemme précédent que $\rho_{jk}^\vartheta(t) = \mathbb{E}[\text{Re}(X_t^j) \text{Im}(X_t^k) | \vartheta]$. Ainsi, en utilisant la proposition 1.13 pour obtenir une forme intégrale de cette espérance conditionnelle, on voit que

$$\rho_{jk}(t) = \int_{\Theta} \mathbb{E}[\text{Re}(X_t^j) \text{Im}(X_t^k) | \vartheta] d\mu(\vartheta) \quad (\text{A.5})$$

$$= \int_{\Theta} \int_{\Omega} \text{Re}(X_t^j(\omega)) \text{Im}(X_t^k(\omega)) d\mathbb{P}^\vartheta(\omega) d\mu(\vartheta) = \mathbb{E}[\text{Re}(X_t^j) \text{Im}(X_t^k)]. \quad (\text{A.6})$$

Les identités pour p_{jk} et r_{jk} sont obtenues de la même façon. ■

On peut maintenant obtenir des expressions pour les dérivées de ces quantités, du moins dans le cas où l'on conditionne sur l'espace des paramètres.

Théorème A.3. Pour tout $\vartheta \in \Theta$, les fonctions p_{jk}^ϑ , r_{jk}^ϑ et ρ_{jk}^ϑ introduites au lemme A.1 sont dérивables sur $(0, \infty)$ et dérивables à droite en 0. De plus, pour tout $t \geq 0$,

$$\begin{aligned} \dot{p}_{jk}^\vartheta(t) &= -\mathbb{E}[\text{Re}(X_t^j) \text{Re}(X_t^k)(b_j(\vartheta, X_t) + b_k(\vartheta, X_t)) | \vartheta] \\ &\quad + \mathbb{E}[\text{Re}(X_t^j)(1 - |X_t^k|)a_k(\vartheta, X_t) | \vartheta] + \mathbb{E}[\text{Re}(X_t^k)(1 - |X_t^j|)a_j(\vartheta, X_t) | \vartheta], \end{aligned} \quad (\text{A.7a})$$

$$\begin{aligned} \dot{r}_{jk}^\vartheta(t) &= -\mathbb{E}[\text{Im}(X_t^j) \text{Im}(X_t^k)(c_j(\vartheta, X_t) + c_k(\vartheta, X_t)) | \vartheta] \\ &\quad + \mathbb{E}[\text{Im}(X_t^j) \text{Re}(X_t^k)b_k(\vartheta, X_t) | \vartheta] + \mathbb{E}[\text{Im}(X_t^k) \text{Re}(X_t^j)b_k(\vartheta, X_t) | \vartheta], \end{aligned} \quad (\text{A.7b})$$

$$\begin{aligned} \dot{\rho}_{jk}^\vartheta(t) &= -\mathbb{E}[\text{Re}(X_t^j) \text{Im}(X_t^k)(b_j(\vartheta, X_t) + c_k(\vartheta, X_t)) | \vartheta] \\ &\quad + \mathbb{E}[\text{Re}(X_t^j) \text{Re}(X_t^k)b_k(\vartheta, X_t) | \vartheta] + \mathbb{E}[\text{Im}(X_t^k)(1 - |X_t^j|)a_j(\vartheta, X_t) | \vartheta]. \end{aligned} \quad (\text{A.7c})$$

Démonstration. Soit $\vartheta \in \Theta$. Comme dans la preuve du théorème 2.6, afin d'alléger la notation, on travaille directement sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P}^\vartheta)$ afin d'éviter le conditionnement sur ϑ . On démontre donc ici l'équivalent des relations (A.7) sur $(\Omega, \mathcal{F}, \mathbb{P}^\vartheta)$.

À partir des résultats démontrés ici, on retrouve le système (A.7) à partir de la relation $\mathbb{E}[Y | \vartheta] = \mathbb{E}^\vartheta[Y]$, qui est vraie pour toute Y mesurable sur (Ω, \mathcal{F}) .

Soient $t \geq 0$ et $\Delta t > 0$. On rappelle que $E = \{0, 1, i\}^N$ est l'ensemble des valeurs possibles de X_t . Ainsi,

$$\rho_{jk}^\vartheta(t + \Delta t) = \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1, X_{t+\Delta t}^k = i] \quad (\text{A.8})$$

$$= \sum_{x \in E} \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1, X_{t+\Delta t}^k = i | X_t = x] \mathbb{P}^\vartheta[X_t = x]. \quad (\text{A.9})$$

Pour $\xi \in \{0, 1, i\}$, on note $f_\xi: \{0, 1, i\} \rightarrow \{0, 1\}$ la fonction qui envoie ξ à 1 et qui envoie les deux autres nombres à zéro. En explicitant les composantes de x dans le conditionnement de la relation précédente, on voit que

$$\begin{aligned} & \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1, X_{t+\Delta t}^k = i | X_t = x] \\ &= \sum_{\xi, \zeta \in \{0, 1, i\}} f_\xi(x_j) f_\zeta(x_k) \mathbb{P}^\vartheta[X_{t+\Delta t}^j = 1, X_{t+\Delta t}^k = i | X_t = x, X_t^j = \xi, X_t^k = \zeta]. \end{aligned} \quad (\text{A.10})$$

D'après la propriétés *iii)* du théorème 2.3, on peut vérifier que lorsque $\Delta t \downarrow 0$, si $x_j = 1$ et $x_k = i$, alors

$$\mathbb{P}[X_{t+\Delta t}^j = 1, X_{t+\Delta t}^k = i | X_t = x] = 1 - b_j(\vartheta, X_t)\Delta t - c_k(\vartheta, X_t)\Delta t + o(\Delta t), \quad (\text{A.11})$$

si $x_j = 1$ et $x_k = 1$ alors

$$\mathbb{P}[X_{t+\Delta t}^j = 1, X_{t+\Delta t}^k = i | X_t = x] = b_k(\vartheta, X_t)\Delta t + o(\Delta t), \quad (\text{A.12})$$

si $x_j = 0$ et $x_k = i$ alors

$$\mathbb{P}[X_{t+\Delta t}^j = 1, X_{t+\Delta t}^k = i | X_t = x] = a_j(\vartheta, X_t)\Delta t + o(\Delta t), \quad (\text{A.13})$$

et si on a à la fois $x_j \neq 1$ et $x_k \neq i$, alors

$$\mathbb{P}[X_{t+\Delta t}^j = 1, X_{t+\Delta t}^k = i | X_t = x] = o(\Delta t). \quad (\text{A.14})$$

En utilisant ces relations, (A.10) devient, lorsque $\Delta \downarrow 0$,

$$\begin{aligned} \mathbb{P}[X_{t+\Delta t}^j = 1, X_{t+\Delta t}^k = i | X_t = x] &= \operatorname{Re}(x_j) \operatorname{Im}(x_k) (1 - b_j(\vartheta, X_t)\Delta t - c_k(\vartheta, X_t)\Delta t) \\ &\quad + \operatorname{Re}(x_j) \operatorname{Re}(x_k) b_k(\vartheta, X_t)\Delta t \\ &\quad + (1 - |x_j|) \operatorname{Im}(x_k) a_j(\vartheta, X_t)\Delta t + o(\Delta t), \end{aligned} \quad (\text{A.15})$$

où l'on a remplacé les fonctions f_ξ et f_ζ par des fonctions appropriées pour chaque cas. En substituant cette relation dans (A.9), on voit que

$$\begin{aligned} \rho_{jk}^\vartheta(t + \Delta t) &= \sum_{x \in E} \left(\operatorname{Re}(x_j) \operatorname{Im}(x_k) (1 - b_j(\vartheta, X_t)\Delta t - c_k(\vartheta, X_t)\Delta t) \right. \\ &\quad + \operatorname{Re}(x_j) \operatorname{Re}(x_k) b_k(\vartheta, X_t)\Delta t \\ &\quad \left. + (1 - |x_j|) \operatorname{Im}(x_k) a_j(\vartheta, X_t)\Delta t \right) \mathbb{P}^\vartheta[X_t = x] + o(\Delta t) \end{aligned} \quad (\text{A.16})$$

Comme E est l'ensemble des valeurs possibles pour X_t , on trouve donc que

$$\begin{aligned}\rho_{jk}^\vartheta(t + \Delta t) &= \mathbb{E}^\vartheta [\operatorname{Re}(X_t^j) \operatorname{Im}(X_t^k) - \operatorname{Re}(X_t^j) \operatorname{Im}(X_t^k)(b_j(\vartheta, X_t) + c_k(\vartheta, X_t))] \Delta t \\ &\quad + \mathbb{E}^\vartheta [\operatorname{Re}(X_t^j) \operatorname{Re}(X_t^k) b_k(\vartheta, X_t)] \Delta t \\ &\quad + \mathbb{E}^\vartheta [(1 - |X_t^j|) \operatorname{Im}(X_t^k) a_j(\vartheta, X_t)] \Delta t + o(\Delta t).\end{aligned}\tag{A.17}$$

Par le lemme A.1, on sait que $\rho_{jk}^\vartheta(t) = \mathbb{E}^\vartheta [\operatorname{Re}(X_t^j) \operatorname{Im}(X_t^k)]$. On voit donc que

$$\begin{aligned}\frac{\rho_{jk}^\vartheta(t + \Delta t) - \rho_{jk}^\vartheta(t)}{\Delta t} &= -\mathbb{E}^\vartheta [\operatorname{Re}(X_t^j) \operatorname{Im}(X_t^k)(b_j(\vartheta, X_t) + c_k(\vartheta, X_t))] \\ &\quad + \mathbb{E}^\vartheta [\operatorname{Re}(X_t^j) \operatorname{Re}(X_t^k) b_k(\vartheta, X_t)] \\ &\quad + \mathbb{E}^\vartheta [(1 - |X_t^j|) \operatorname{Im}(X_t^k) a_j(\vartheta, X_t)] + \frac{1}{\Delta t} o(\Delta t).\end{aligned}\tag{A.18}$$

Puisque $\frac{1}{\Delta t} o(\Delta t) \rightarrow 0$ lorsque $\Delta t \downarrow 0$, la relation précédente montre donc que la dérivée à droite de ρ_{jk}^ϑ existe et est donnée par (A.7c). Comme l'équation de Kolmogorov montre que ρ_{jk}^ϑ est dérivable sur $(0, \infty)$, ce résultat donne la valeur de la dérivée $\dot{\rho}_{jk}^\vartheta(t)$.

Un raisonnement semblable mène aux expressions des dérivées de p_{jk}^ϑ et de r_{jk}^ϑ . ■

On obtient ensuite le corollaire suivant où l'on élimine le conditionnement sur l'espace des paramètres.

Corollaire. Si pour tout $x \in E$ et pour tout nœud j de G , les fonctions $a_j(\cdot, x)$, $b_j(\cdot, x)$ et $c_j(\cdot, x)$ sont intégrables, alors les fonctions p_{jk} , r_{jk} et ρ_{jk} introduites au lemme A.2 sont dérивables sur $(0, \infty)$ et dérивables à droite en 0, et pour tout $t \geq 0$,

$$\begin{aligned}\dot{p}_{jk}(t) &= -\mathbb{E} [\operatorname{Re}(X_t^j) \operatorname{Re}(X_t^k)(b_j(\vartheta, X_t) + b_k(\vartheta, X_t))] \\ &\quad + \mathbb{E} [\operatorname{Re}(X_t^j)(1 - |X_t^k|) a_k(\vartheta, X_t)] + \mathbb{E} [\operatorname{Re}(X_t^k)(1 - |X_t^j|) a_j(\vartheta, X_t)],\end{aligned}\tag{A.19a}$$

$$\begin{aligned}\dot{r}_{jk}(t) &= -\mathbb{E} [\operatorname{Im}(X_t^j) \operatorname{Im}(X_t^k)(c_j(\vartheta, X_t) + c_k(\vartheta, X_t))] \\ &\quad + \mathbb{E} [\operatorname{Im}(X_t^j) \operatorname{Re}(X_t^k) b_k(\vartheta, X_t)] + \mathbb{E} [\operatorname{Im}(X_t^k) \operatorname{Re}(X_t^j) b_k(\vartheta, X_t)],\end{aligned}\tag{A.19b}$$

$$\begin{aligned}\dot{\rho}_{jk}(t) &= -\mathbb{E} [\operatorname{Re}(X_t^j) \operatorname{Im}(X_t^k)(b_j(\vartheta, X_t) + c_k(\vartheta, X_t))] \\ &\quad + \mathbb{E} [\operatorname{Re}(X_t^j) \operatorname{Re}(X_t^k) b_k(\vartheta, X_t)] + \mathbb{E} [\operatorname{Im}(X_t^k)(1 - |X_t^j|) a_j(\vartheta, X_t)].\end{aligned}\tag{A.19c} \blacklozenge$$

Démonstration. On note d'abord que pour tout $\vartheta \in \Theta$, pour tout $t \geq 0$ et pour chaque paire de nœuds j, k du graphe, le théorème A.3 montre que

$$\begin{aligned}|\rho_{jk}^\vartheta(t)| &\leq \mathbb{E} [|\operatorname{Re}(X_t^j) \operatorname{Im}(X_t^k)(b_j(\vartheta, X_t) + c_k(\vartheta, X_t))| \mid \vartheta] \\ &\quad + \mathbb{E} [|\operatorname{Re}(X_t^j) \operatorname{Re}(X_t^k) b_k(\vartheta, X_t)| \mid \vartheta] + \mathbb{E} [|\operatorname{Im}(X_t^k)(1 - |X_t^j|) a_j(\vartheta, X_t)| \mid \vartheta]\end{aligned}\tag{A.20}$$

$$\leq \mathbb{E} [|b_j(\vartheta, X_t)| + |c_k(\vartheta, X_t)| + |b_k(\vartheta, X_t)| + |a_j(\vartheta, X_t)| \mid \vartheta].\tag{A.21}$$

Comme E est fini, on peut définir

$$f(\vartheta) := \sum_{x \in E} (|b_j(\vartheta, X_t)| + |c_k(\vartheta, X_t)| + |b_k(\vartheta, X_t)| + |a_j(\vartheta, X_t)|), \quad (\text{A.22})$$

qui est intégrable puisque l'on suppose que les fonctions $a_\ell(\cdot, x)$, $b_\ell(\cdot, x)$ et $c_\ell(\cdot, x)$ sont intégrables pour chaque $x \in E$. Cette fonction permet de borner $|\rho_{jk}^\vartheta(t)|$:

$$|\rho_{jk}^\vartheta(t)| \leq \sum_{x \in E} \mathbb{E}[|b_j(\vartheta, X_t)| + |c_k(\vartheta, X_t)| + |b_k(\vartheta, X_t)| + |a_j(\vartheta, X_t)| \mid \vartheta] = \mathbb{E}[f(\vartheta) \mid \vartheta], \quad (\text{A.23})$$

de sorte que $|\rho_{jk}^\vartheta(t)| \leq f(\vartheta)$. Comme f est intégrable, cette condition de domination justifie de passer la dérivée sous l'intégrale, et on trouve que

$$\dot{\rho}_{jk}(t) = \frac{d}{dt} \int_{\Theta} \rho_{jk}^\vartheta(t) d\mu(\vartheta) = \int_{\Theta} \dot{\rho}_{jk}^\vartheta(t) d\mu(\vartheta). \quad (\text{A.24})$$

L'expression de la dérivée de ρ_{jk} suit alors directement de l'intégration de l'expression de ρ_{jk}^ϑ obtenue au théorème A.3, en utilisant la proposition 1.13. ■

Les deux autres résultats peuvent être démontrés de la même façon.

On veut ensuite passer aux résultats analogues sur les populations de graphe. Pour ce faire, on procède de la même façon qu'à la section 2.4. On suppose que le graphe est divisé en une collection \mathcal{P} de n populations. Afin d'alléger la notation, on note $E_{UV}^{JK}(t) := \mathbb{E}[U_t^J V_t^K]$, où U et V remplacent A , R ou S . Alors

$$E_{AR}^{JK}(t) = \mathbb{E}[A_t^J A_t^K] = \mathbb{E}\left[\frac{1}{|J||K|} \sum_{j \in J, k \in K} \operatorname{Re}(X_t^j) \operatorname{Im}(X_t^k)\right] = \frac{1}{|J||K|} \sum_{j \in J, k \in K} \rho_{jk}(t), \quad (\text{A.25})$$

de sorte que

$$\dot{E}_{AR}^{JK}(t) = \frac{1}{|J||K|} \sum_{j \in J, k \in K} \dot{\rho}_{jk}(t). \quad (\text{A.26})$$

Des identités semblables peuvent évidemment être obtenues dans les autres cas.

À partir de (A.26) et de l'expression de $\dot{\rho}_{jk}^\vartheta$ donnée au corollaire du théorème A.3, on calcule que

$$\begin{aligned} \dot{E}_{AR}^{JK}(t) &= \frac{1}{|J||K|} \sum_{j \in J, k \in K} \left(-\mathbb{E}[\operatorname{Re}(X_t^j) \operatorname{Im}(X_t^k) (b_j(\vartheta, X_t) + c_k(\vartheta, X_t))] \right. \\ &\quad + \mathbb{E}[\operatorname{Re}(X_t^j) \operatorname{Re}(X_t^k) b_k(\vartheta, X_t)] \\ &\quad \left. + \mathbb{E}[\operatorname{Im}(X_t^k) (1 - |X_t^j|) a_j(\vartheta, X_t)] \right). \end{aligned} \quad (\text{A.27})$$

Comme l'espérance est linéaire, il s'ensuit que

$$\begin{aligned}
 \dot{\mathbb{E}}_{AR}^{JK}(t) = & -\mathbb{E}\left[\frac{1}{|J|} \sum_{j \in J} \operatorname{Re}(X_t^j) b_j(\vartheta, X_t) \frac{1}{|K|} \sum_{k \in K} \operatorname{Im}(X_t^k)\right] \\
 & -\mathbb{E}\left[\frac{1}{|J|} \sum_{j \in J} \operatorname{Re}(X_t^j) \frac{1}{|K|} \sum_{k \in K} \operatorname{Im}(X_t^k) c_k(\vartheta, X_t)\right] \\
 & +\mathbb{E}\left[\frac{1}{|J|} \sum_{j \in J} \operatorname{Re}(X_t^j) \frac{1}{|K|} \sum_{k \in K} \operatorname{Re}(X_t^k) b_k(\vartheta, X_t)\right] \\
 & +\mathbb{E}\left[\frac{1}{|J|} \sum_{j \in J} (1 - |X_t^j|) a_j(\vartheta, X_t) \frac{1}{|K|} \sum_{k \in K} \operatorname{Im}(X_t^k)\right].
 \end{aligned} \tag{A.28}$$

On réintroduit alors les ensembles $J_t^\xi = \{j \in J : X_t^j = \xi\}$ pour $\xi \in \{0, 1, i\}$. On note que $A_t^J = |J_t^1|/|J|$, que $R_t^J = |J_t^i|/|J|$ et que $S_t^J = |J_t^0|/|J|$. Ainsi,

$$\begin{aligned}
 \dot{\mathbb{E}}_{AR}^{JK}(t) = & -\mathbb{E}\left[\frac{A_t^J}{|J_t^1|} \sum_{j \in J_t^1} b_j(\vartheta, X_t) R_t^K\right] - \mathbb{E}\left[A_t^J \frac{R_t^K}{|K_t^i|} \sum_{k \in K_t^i} c_k(\vartheta, X_t)\right] \\
 & +\mathbb{E}\left[A_t^J \frac{A_t^K}{|K_t^1|} \sum_{k \in K_t^1} b_k(\vartheta, X_t)\right] + \mathbb{E}\left[\frac{S_t^J}{|J_t^0|} \sum_{j \in J_t^0} a_j(\vartheta, X_t) R_t^K\right].
 \end{aligned} \tag{A.29}$$

Afin d'obtenir les dérivées des seconds moments, on suppose comme à la section 2.4 que pour chaque population $J \in \mathcal{P}$, il existe des fonctions $a_J, b_J, c_J : [0, 1]^{2n} \rightarrow [0, \infty)$ telles que

$$a_J(Y_t) = \frac{1}{|J_t^0|} \sum_{j \in J_t^0} a_j(\vartheta, X_t), \quad b_J(Y_t) = \frac{1}{|J_t^1|} \sum_{j \in J_t^1} b_j(\vartheta, X_t) \quad \text{et} \quad c_J(Y_t) = \frac{1}{|J_t^i|} \sum_{j \in J_t^i} c_j(\vartheta, X_t), \tag{A.30}$$

où les égalités tiennent \mathbb{P} -p.s. et où Y_t est l'état macroscopique qui correspond à l'état microscopique X_t selon (2.49). Si de telles fonctions existent, alors (A.29) devient

$$\dot{\mathbb{E}}_{AR}^{JK}(t) = -\mathbb{E}[A_t^J R_t^K (b_J(Y_t) + c_K(Y_t))] + \mathbb{E}[A_t^J A_t^K b_K(Y_t)] + \mathbb{E}[S_t^J R_t^K a_J(Y_t)]. \tag{A.31}$$

De la même façon, on peut montrer que

$$\dot{\mathbb{E}}_{AA}^{JK}(t) = -\mathbb{E}[A_t^J A_t^K (b_J(Y_t) + b_K(Y_t))] + \mathbb{E}[A_t^J S_t^K a_K(Y_t)] + \mathbb{E}[S_t^J A_t^K a_J(Y_t)] \tag{A.32}$$

et que

$$\dot{\mathbb{E}}_{RR}^{JK}(t) = -\mathbb{E}[R_t^J R_t^K (c_J(Y_t) + c_K(Y_t))] + \mathbb{E}[R_t^J A_t^K b_K(Y_t)] + \mathbb{E}[A_t^J R_t^K b_J(Y_t)]. \tag{A.33}$$

Les dérivées des seconds moments qui impliquent des fractions de populations dans l'état 0 peuvent alors être obtenues à partir du fait que $A_t^J + R_t^J + S_t^J = 1$ pour tout $t \geq 0$, pour toute population J . En effet, cette relation montre par exemple que $\mathbb{E}_{AS}^{JK} = \mathcal{A}_J - \mathbb{E}_{AA}^{JK} - \mathbb{E}_{AR}^{JK}$.

Ainsi, $\dot{\mathbb{E}}_{AS}^{JK} = \dot{\mathcal{A}}_J - \dot{\mathbb{E}}_{AA}^{JK} - \dot{\mathbb{E}}_{AR}^{JK}$. D'après les résultats (A.31) et (A.32) et d'après l'expression (2.60a) de la dérivée de \mathcal{A}_J , on a alors que

$$\begin{aligned}\dot{\mathbb{E}}_{AS}^{JK}(t) &= -\mathbb{E}[A_t^J b_J(Y_t)] + \mathbb{E}[S_t^J a_J(Y_t)] \\ &\quad + \mathbb{E}[A_t^J A_t^K (b_J(Y_t) + b_K(Y_t))] - \mathbb{E}[A_t^J S_t^K a_K(Y_t)] - \mathbb{E}[S_t^J A_t^K a_J(Y_t)] \quad (\text{A.34})\end{aligned}$$

$$\begin{aligned}&\quad + \mathbb{E}[A_t^J A_t^K (b_J(Y_t) + c_K(Y_t))] - \mathbb{E}[A_t^J A_t^K b_K(Y_t)] - \mathbb{E}[S_t^J R_t^K a_J(Y_t)] \\ &= -\mathbb{E}[A_t^J S_t^K a_K(Y_t)] - \mathbb{E}[A_t^J (1 - A_t^K - R_t^K) b_J(Y_t)] \quad (\text{A.35}) \\ &\quad + \mathbb{E}[A_t^J R_t^K c_K(Y_t)] + \mathbb{E}[S_t^J (1 - A_t^K - R_t^K) a_J(Y_t)]\end{aligned}$$

$$= -\mathbb{E}[A_t^J S_t^K (b_J(Y_t) + a_K(Y_t))] + \mathbb{E}[A_t^J R_t^K c_K(Y_t)] + \mathbb{E}[S_t^J S_t^K a_J(Y_t)]. \quad (\text{A.36})$$

On peut obtenir de la même façon que

$$\dot{\mathbb{E}}_{RS}^{JK}(t) = -\mathbb{E}[R_t^J S_t^K (c_J(Y_t) + a_K(Y_t))] + \mathbb{E}[R_t^J R_t^K c_K(Y_t)] + \mathbb{E}[A_t^J S_t^K b_J(Y_t)] \quad (\text{A.37})$$

et que

$$\dot{\mathbb{E}}_{SS}^{JK}(t) = -\mathbb{E}[S_t^J S_t^K (a_J(Y_t) + a_K(Y_t))] + \mathbb{E}[S_t^J R_t^K c_K(Y_t)] + \mathbb{E}[R_t^J S_t^K c_J(Y_t)]. \quad (\text{A.38})$$

Ceci termine de dériver le système des seconds moments non centrés (2.61).

Les résultats pour les dérivées des produits mènent directement au système des seconds moments centrés (2.62). Pour alléger, on note $C_{UV}^{JK} := \text{Cov}[U_t^J, V_t^K]$ où U et V remplacent A , R ou S . Alors $C_{AR}^{JK} = \mathbb{E}_{AR}^{JK} - \mathcal{A}_J \mathcal{R}_K$, de sorte que

$$\dot{C}_{AR}^{JK}(t) = \dot{\mathbb{E}}_{AR}^{JK} - \dot{\mathcal{A}}_J \mathcal{R}_K(t) - \mathcal{A}_J(t) \dot{\mathcal{R}}_K. \quad (\text{A.39})$$

En utilisant les résultats (A.31) et (2.60), on calcule que

$$\begin{aligned}\dot{C}_{AR}^{JK}(t) &= -\mathbb{E}[A_t^J R_t^K (b_J(Y_t) + c_K(Y_t))] + \mathbb{E}[A_t^J A_t^K b_K(Y_t)] + \mathbb{E}[S_t^J R_t^K a_J(Y_t)] \\ &\quad + \mathbb{E}[A_t^J b_J(Y_t)] \mathcal{R}_K(t) - \mathbb{E}[S_t^J a_J(Y_t)] \mathcal{R}_K(t) \quad (\text{A.40}) \\ &\quad + \mathcal{A}_J(t) \mathbb{E}[R_t^K c_K(Y_t)] - \mathcal{A}_J(t) \mathbb{E}[A_t^K b_K(Y_t)]\end{aligned}$$

$$\begin{aligned}&= -\text{Cov}[A_t^J b_J(Y_t), R_t^K] - \text{Cov}[A_t^J, R_t^K c_K(Y_t)] \quad (\text{A.41}) \\ &\quad + \text{Cov}[A_t^J, A_t^K b_K(Y_t)] + \text{Cov}[S_t^J a_J(Y_t), R_t^K].\end{aligned}$$

De la même façon, on peut obtenir que

$$\begin{aligned}\dot{C}_{AA}^{JK}(t) &= -\text{Cov}[A_t^J, A_t^K b_K(Y_t)] - \text{Cov}[A_t^J b_J(Y_t), A_t^K] \\ &\quad + \text{Cov}[A_t^J, S_t^K a_K(Y_t)] + \text{Cov}[S_t^J a_J(Y_t), A_t^K], \quad (\text{A.42})\end{aligned}$$

et que

$$\begin{aligned}\dot{C}_{RR}^{JK}(t) &= -\text{Cov}[R_t^J, R_t^K c_K(Y_t)] - \text{Cov}[R_t^J c_J(Y_t), R_t^K] \\ &\quad + \text{Cov}[R_t^J, A_t^K b_K(Y_t)] + \text{Cov}[A_t^J b_J(Y_t), R_t^K], \quad (\text{A.43})\end{aligned}$$

ce qui termine de dériver le système (2.62).

Annexe B

PopNet

Aux chapitres 3 et 4, plusieurs résultats d'intégrations numériques d'équations différentielles et de simulations de processus stochastiques ont été présentés. Tous ces résultats ont été obtenus à partir du paquetage PopNet [77], écrit dans le langage Python. On présente ici brièvement les idées importantes des méthodes utilisées pour réaliser ces expériences numériques. Pour plus de détails, on réfère à la documentation de PopNet.

§ B.1. À propos des intégrations numériques

Les intégrations numériques d'équations différentielles dont des résultats sont présentés dans ce mémoire ont toutes été réalisées à l'aide de la classe `ode` du module `integrate` de la librairie SciPy [96]. Cette classe fournit plusieurs intégrateurs numériques, mais l'intégrateur `vode` a été utilisé dans tous les cas présentés ici.

§ B.2. À propos des simulations du processus stochastique

PopNet permet également de simuler numériquement des trajectoires du processus stochastique présenté à la section 4.2, qui décrit l'évolution de l'état d'un réseau de neurones biologiques d'un point de vue microscopique. Ces simulations sont faites à partir de l'algorithme de Doob–Gillespie [37]. Plus de détails sont donnés dans la documentation de la classe `Simulator`, dans le module `executors` de PopNet.

Pour toutes les simulations dont des résultats sont présentés au chapitre 4, les valeurs des paramètres α , β et γ sont fixées à leurs valeurs moyennes pour tous les neurones. De même, la valeur d'un poids de connexion W_{jk} entre deux neurones k et j est toujours fixée à la valeur prévue par le coefficient c_{JK} qui décrit la connexion de la population $K \ni k$ vers la population $J \ni j$ selon la relation $W_{jk} = W_{JK} = c_{JK}/|K|$, où $|K|$ est la taille de K . Toutefois, les valeurs des seuils d'activation θ sont tirées de lois logistiques dont les paramètres sont spécifiés dans chaque cas. Finalement, l'état initial X_0^j d'un neurone j d'une population J est choisi aléatoirement selon une loi de probabilité telle que $\mathbb{P}[X_0^j = 1] = \mathcal{A}_{J0}$, $\mathbb{P}[X_0^j = i] = \mathcal{R}_{J0}$ et $\mathbb{P}[X_0^j = 0] = \mathcal{S}_{J0}$, où \mathcal{A}_{J0} , \mathcal{R}_{J0} et \mathcal{S}_{J0} sont les espérances initiales des fractions de J dans chaque état.