



HAL
open science

Simulation, estimation paramétrique et classification des trajectoires d'un processus birth-death-move avec mutations.

Lisa Balsollier

► **To cite this version:**

Lisa Balsollier. Simulation, estimation paramétrique et classification des trajectoires d'un processus birth-death-move avec mutations.. Mathématiques [math]. Nantes université, 2024. Français. NNT : . tel-04881209

HAL Id: tel-04881209

<https://hal.science/tel-04881209v1>

Submitted on 22 Jan 2025

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

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

THÈSE DE DOCTORAT DE

NANTES UNIVERSITÉ

ÉCOLE DOCTORALE N° 641
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*

Spécialité : *Mathématiques et leurs interactions*

Par

Lisa BALSOLLIER

**Simulation, estimation paramétrique et classification des
trajectoires d'un processus birth-death-move avec mutations.**

Thèse présentée et soutenue au Centre INRIA de l'Université de Rennes, le 21 octobre 2024
Unité de recherche : Laboratoire de Mathématiques Jean Leray

Rapporteurs avant soutenance :

Édith GABRIEL Directrice de Recherche, INRAE
Thibault LAGACHE Chargé de Recherche, Institut Pasteur, Paris

Composition du Jury :

Examineurs :	Marie-Pierre ÉTIENNE	Professeure associée, ENSAI, Rennes
	Ronan LE GUÉVEL	Maître de conférence, Université de Rennes
Dir. de thèse :	Frédéric LAVANCIER	Professeur, ENSAI, Rennes
Co-dir. de thèse :	Charles KERVRANN	Directeur de Recherche, Centre INRIA de l'Université de Rennes
	Jean SALAMERO	Directeur de Recherche, Institut de Génétique Humaine, Montpellier

REMERCIEMENTS

L'accomplissement de cette thèse a été une étape importante de mon cursus universitaire et son achèvement marque la fin de mon parcours scolaire. À cette occasion, je formule avec sincérité ma gratitude envers tous ceux qui m'ont accompagnée tout au long de ce chemin.

Je tiens, tout d'abord, à témoigner ma profonde reconnaissance envers mes directeurs de thèse, Frédéric Lavancier et Charles Kervrann pour m'avoir accueillie pour mon stage de Master, puis pour avoir accepté de m'encadrer pour cette thèse. Je les remercie pour leurs conseils, leur patience et nos échanges constructifs qui m'ont permis de surmonter les difficultés et de mener à bien mon travail.

Je voudrais également remercier Jean Salamero qui a suivi l'avancée de mes travaux avec son regard de biologiste, me permettant de découvrir un autre domaine de la recherche.

Je remercie Édith Gabriel et Thibault Lagache de m'avoir fait l'honneur de rapporter mon mémoire de thèse et également Marie-Pierre Étienne et Ronan Le Guével d'avoir accepté de faire partie de mon jury de thèse. Je souhaite aussi mentionner mon comité de suivi de thèse avec Lüdger Johannes, aux côtés de Ronan Le Guével, dont les encouragements et les conseils m'ont été d'une grande aide.

J'aimerais exprimer toute ma gratitude envers les équipes administratives de l'Université de Nantes, de Rennes et du Centre INRIA, et plus particulièrement Béatrice Havet, Marie-Aude Verger et Caroline Tanguy pour leur assistance précieuse tout au long de mon parcours doctoral. Leur professionnalisme et leur disponibilité ont allégé mes démarches et facilité ainsi mon travail.

Je remercie mes camarades doctorants (ou ex-doctorants) qui ont partagé avec moi cette période intense de recherche, notamment Sébastien et Yasmine qui ont été de formidables collègues de bureau.

J'ai aussi une pensée très affectueuse pour Fabrice, Emeline et Laura, qui ont toujours veillé sur moi, comme des aînés, durant ces 7 dernières années. Leur soutien

et leurs conseils avisés ont été une véritable source de motivation. Merci infiniment !

Pendant ces trois années, j'ai eu le privilège d'intervenir dans la préparation à l'agrégation interne de mathématiques de l'Université de Rennes 1 et je remercie chaleureusement David Bourqui, Bernard Delyon et Stéphane Le Borgne de m'avoir fait confiance pour accomplir cette mission. J'ai pris un réel plaisir à partager mes connaissances avec ces étudiantes et étudiants « de luxe » dans une ambiance studieuse et stimulante.

Parallèlement, mes interventions en licence de biologie et d'informatique ont été une expérience bénéfique et enrichissante pour moi.

Je voudrais aussi exprimer ma reconnaissance envers tous les professeurs de mathématiques qui ont marqué ma scolarité. Grâce à eux les mathématiques sont devenues pour moi, bien plus qu'une matière parmi d'autres. Du collège avec Mesdames Crumière et Buratti, à mes années de prépas avec Messieurs Zannad et Arzac, en passant par mes années de lycée avec Monsieur Vidal et Mesdames Tissier et Pommier, j'ai rencontré des enseignants remarquables, exigeants et investis. Leur dévouement et la qualité de leur enseignement resteront des sources d'inspiration pour moi. Merci du fond du cœur d'avoir nourri et cultivé ma passion des mathématiques.

Même si son tour viendra plus loin dans ces remerciements, je ne peux pas terminer ce paragraphe sur mes profs de maths sans parler de celle qui a été la meilleure professeure : merci mille fois, Maman, pour tout le temps que tu as passé à rédiger des textes simplifiés sur des notions compliquées afin de rendre plus intuitifs des concepts abstraits, ton aide me faisait gagner un temps tellement précieux dans les années difficiles qu'ont été la prépa et l'agreg.

Mon parcours aurait été sans doute plus monotone sans la complicité de mes amis de toujours : les Bestous et plus particulièrement Lucile, Luna, Mathilde et Carla. Malgré notre éloignement quotidien, vous restez des piliers de ma vie. Lucile, merci pour les heures que tu as passées à m'écouter et tous les mots justes que tu as su trouver pour me remonter le moral. Merci pour tout ce que l'on a vécu et pour tout ce que l'on va encore vivre ensemble.

Je pense aussi aux amis que j'ai rencontrés lors de mon arrivée à Rennes, il y a 7 ans maintenant. Tout d'abord, merci à ma quasi-coloc Angèle pour notre entre-aide, notre complicité et nos rires, aussi bien à l'école que dans notre vie quotidienne.

Je me souviens aussi de notre fameux trio avec Briec.

Merci aussi à l'ABP (la vraie !) avec Kim, Mahé, Kaber et Adèle, ainsi qu'à Arthur, Olivier, Aymeric, Julien, Caj. Votre amitié est irremplaçable pour moi. Vous êtes une équipe formidable, et je suis vraiment heureuse d'avoir partagé avec vous tant de bons moments, qui resteront au fond de mon cœur pour toujours. Bien sûr, « on s'arrête pas là, jusqu'à où est-ce qu'on ira ? »

Cette thèse n'aurait jamais pu aboutir sans les nombreuses séances au club d'aviation qui, dans un premier temps, me permettaient de me vider la tête. J'y retrouvais avec plaisir la bonne humeur de Guillaume, Pauline, Régis, Romain et Benoît. Maintenant, même si la page se tourne, ce sport restera une véritable source d'équilibre pour moi.

Je mesure combien je suis chanceuse d'avoir une famille si aimante, toujours présente pour m'accompagner et toujours prête pour m'aider dans n'importe quelle circonstance.

Merci Enzo pour ton optimisme et tes encouragements qui m'ont souvent permis de voir les choses sous un autre angle. Nos différences unies par un amour fraternel sont une force commune.

Merci Papa et Maman pour votre présence et votre dévouement inestimable.

Merci Papa pour ton pragmatisme, ta gentillesse et ta sérénité en toute circonstance. Merci Maman pour ton écoute, ta dévotion désintéressée, ta détermination hors du commun pour résoudre n'importe quel problème, qu'il soit mathématique ou existentiel.

Merci ma Mémé chérie, toujours là pour moi sans faillir, tu vois toujours le bon côté des choses et tu m'encourages comme personne, chaque matin. Tu as toujours cru en moi bien plus que moi-même et je vais enfin pouvoir te dire « j'ai fini l'école ! ». Je pense aussi à Pépé et Mémé Maurise qui auraient été si heureux...

Enfin et surtout, un immense merci à toi, Émilien. Je n'aurais jamais pu imaginer un soutien plus précieux, un confident plus bienveillant, un complice plus fidèle, un coéquipier plus volontaire, un allié plus discret, pour franchir les étapes de ce long chemin que nous parcourons ensemble. Derrière chacune de mes réussites, il y a une part de toi.

SOMMAIRE

Introduction	11
1 Mise en contexte et objectifs de la thèse	11
1.1 Modéliser une dynamique moléculaire	11
1.2 Le processus birth-death-move avec mutations	13
1.3 État de l'art et contributions de la thèse	16
2 Organisation de la thèse	18
I A generative model to simulate the spatio-temporal dynamics of biomolecules in cells.	25
1 Presentation of the birth-death-move process with mutations	27
1.1 Heuristic and notations	27
1.2 Algorithmic construction	29
1.3 Exemplified specifications of the model	31
1.3.1 The inter-jumps motion	31
1.3.2 The intensity functions	33
1.3.3 The transition probability functions	35
1.4 Conclusion	38
2 Application to the study of the dynamics of real data	39
2.1 Application to the joint dynamics of Langerin and Rab-11 proteins .	39
2.1.1 Description of the dataset	39
2.1.2 Simulation of synthetic sequences	44
2.2 Conclusion and perspectives	47
Appendix of Part I	53
A Simulation algorithm	53

II Parametric estimation and LAN property of a BDMM process.	57
3 Likelihood and LAN property of a BDMM process.	59
3.1 Definition of the BDMM process	59
3.1.1 State space	59
3.1.2 Dynamics of a BDMM process	60
3.1.3 Elements of the model and examples	63
3.2 LAN property	69
3.2.1 Likelihood of the BDMM process	69
3.2.2 LAN	70
3.2.3 Examples	74
3.3 Conclusion	76
4 Analysis of the colocation phenomenon on simulated and real data.	77
4.1 Simulation study	77
4.2 Data analysis	80
4.3 Conclusion	82
Appendix of Part II	85
B Remark about the diffusion coefficient in case of continuous time observations	85
C Proof of the likelihood expression of a BDMM process	87
D Proof of the LAN property of a BDMM process	93
E Ergodicity of the BDMM process	107
III Analyse du mouvement des trajectoires et détection des instants de changement de régime.	111
5 Classification des trajectoires	113
5.1 État de l'art	114
5.2 Résolution du problème	116

5.2.1	Notations	116
5.2.2	Algorithme EM pour estimer les paramètres qui maximisent la vraisemblance de la séquence observée	119
5.2.3	Algorithme de Viterbi pour estimer les états cachés ayant pro- duit la séquence observée	126
5.3	Étude théorique d'un exemple dans le cas de trois mouvements	130
5.3.1	Présentation de l'exemple	130
5.3.2	Un argument du maximum de la fonction Q explicite dans cet exemple	135
5.4	Conclusion	142
6	Applications sur des données simulées et réelles	143
6.1	Résultats sur des données simulées	143
6.1.1	Cas où seuls les paramètres de la matrice de transition a sont inconnus	146
6.1.2	Cas où les paramètres de chaque trajectoire sont inconnus et distincts	147
6.1.3	Cas où les paramètres $\sigma_{i,1}, \rho_i, \sigma_{i,2}, \lambda_i, \sigma_{i,3}$ sont identiques pour toutes les trajectoires	152
6.1.4	Bilan de l'étude par simulation	157
6.2	Résultats sur des données réelles	159
6.3	Conclusion	161
	Annexes de la partie III	164
F	Forme de la fonction Q dans le cas de trajectoires indépendantes	165
G	Preuves du calcul de l'argument du maximum de Q	199
H	Comparaison avec la méthode de [MBP⁺15]	233
	Conclusion	241
	Bibliographie	247

INTRODUCTION

1 Mise en contexte et objectifs de la thèse

1.1 Modéliser une dynamique moléculaire

Le déplacement des particules qui interagissent dans les cellules ainsi que leur organisation dans le temps et dans l'espace est un sujet de questionnement majeur et permanent en biologie cellulaire. La compréhension de ces mouvements est un objectif à long terme de la recherche scientifique qui mobilise des physiciens, des biologistes, des informaticiens et des mathématiciens.

Pour expliquer ces phénomènes, des méthodes d'analyse d'images ont été développées afin de traiter des vidéos acquises en microscopie de fluorescence [SFS⁺97, GCK⁺18, AABA10, BKB09, Bou23, LCB⁺20]. Ces études demandent dans un premier temps un suivi précis (tracking) des particules considérées puis dans un deuxième temps une méthode fiable pour interpréter le mouvement des particules (e.g. protéines) en question. Dans cette thèse, nous ne nous intéresserons pas aux méthodes utilisées pour effectuer la tâche d'estimation des trajectoires, qui sont déjà très nombreuses [KSA⁺15, CSDC⁺14, RDJ⁺17, GFMS20, VVP14, VHHK16, SPB⁺19, LCB⁺20, BAD⁺21, Bou23], mais nous nous consacrerons à l'analyse de la dynamique des particules (apparition, mouvement, disparition) à partir des trajectoires estimées.

Afin d'exemplifier le type de dynamique qui nous intéresse, la figure 1 illustre (à gauche) une cellule vivante extraite d'une séquence vidéo, acquise en microscopie TIRF (total internal reflection fluorescence) par l'équipe de Jean Salamero de l'Institut Curie (UMR 144 CNRS). Les petites taches (ou spots) visibles témoignent de la présence de protéines de type Langerin, à proximité de la membrane plasmique de la cellule. À droite de la figure 1, nous affichons les trajectoires extraites de cette séquence après traitement informatique (segmentation, filtrage stochastique).

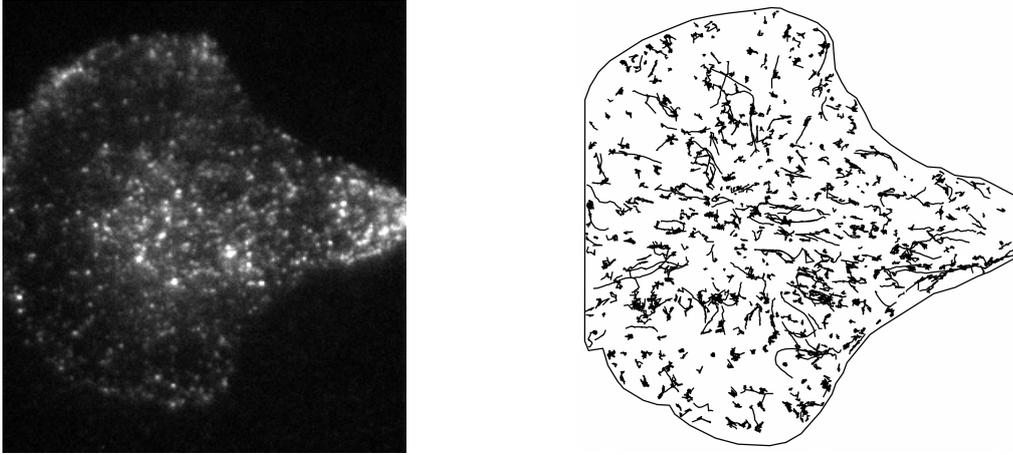


FIGURE 1 – Dynamique des protéines de type Langerin près de la membrane d’une cellule lors du phénomène d’exocytose. À gauche : une image de la séquence vidéo brute obtenue par microscopie TIRF. À droite : superposition de toutes les trajectoires obtenues après traitement de l’ensemble de la vidéo brute (comportant 1200 images).

Les modèles mathématiques et biophysiques sont indispensables pour simuler et analyser le comportement des molécules dans les cellules. Les données simulées permettent non seulement d’évaluer les performances des algorithmes d’analyse d’images, mais aussi d’entraîner des modèles complexes comme les réseaux neuronaux profonds. Dans ce contexte, les modèles stochastiques basés sur des particules sont largement utilisés pour le suivi des molécules [CSDC⁺14, RDJ⁺17, GFMS20, VVP14] et servent aussi à créer des simulateurs de microscopie de localisation de molécules uniques (SMLM) [VHHK16, SPB⁺19, LCB⁺20, BAD⁺21, Bou23].

Au-delà des logiciels usuels tels que Virtual Cell [SFS⁺97], MCell [GCK⁺18] et Smoldyn [AABA10], spécialisés dans la caractérisation des dynamiques réaction-diffusion, de nombreux modèles stochastiques ont été mis en place afin de représenter le comportement des particules dans la cellule, en incluant des mouvements browniens, confinés, anormaux ou dirigés avec des vitesses variables [BKB09, Bou23, BN13, HH17, BKV18, PZB⁺18, BVK19, LCB⁺20, ASP⁺19, MGVGGM⁺21], pour analyser le comportement des molécules dans le cytoplasme ou lorsqu’elles sont propulsées le long d’un réseau de cytosquelette [BKB09, LDH09, Ker22]

Cependant, toutes ces dynamiques ne prennent pas en compte des changements de régime au cours du déplacement des particules, bien que ce comportement soit manifestement observé dans des applications réelles [BKV18]. De plus, ces modèles

supposent généralement que les particules sont indépendantes et considèrent un taux de naissance constant. Ils ne tiennent pas non plus compte de l'influence des particules voisines sur le moment et le lieu d'apparition de nouvelles particules, en excluant toute interaction avec les particules existantes. Les mêmes limites sont présentes si on s'intéresse à la dynamique de disparition des particules. Il existe donc encore des opportunités pour améliorer les modèles existants afin de prendre en considération ces spécificités de déplacements des molécules.

1.2 Le processus birth-death-move avec mutations

Dans cette thèse, nous proposons de tirer parti d'une adaptation d'un modèle stochastique, appelé processus "birth-death-move" (BDM), qui est suffisamment flexible pour inclure toutes les fonctionnalités mentionnées précédemment dans un cadre unifié et théoriquement bien fondé.

Les processus BDM [LLG21] ont été introduits pour enrichir la notion déjà formalisée par Preston depuis 1975 [Pre75] des processus de naissance-mort spatiaux (dits spatial birth-death processes). Ces derniers ont eux-mêmes permis d'élargir la notion de processus de naissance-mort simple, qui ne considérait que le cas où les paramètres du système dépendent seulement de la taille de la population d'individus ou de particules. La vision de Preston permet de considérer des cas beaucoup plus complets, dans lesquels les apparitions et disparitions des individus (particules dans notre cas) dépendent de la position des autres individus. On peut ainsi traiter des modèles beaucoup plus généraux où, par exemple, les individus ont tendance à apparaître à proximité d'individus déjà vivants (phénomènes de propagation), ou bien à disparaître s'ils sont trop nombreux dans un même site (phénomènes de concurrence des ressources). Néanmoins, une limite du modèle de Preston est qu'il suppose que la localisation spatiale de chaque individu reste fixe, de sa naissance jusqu'à sa mort. C'est pour palier cette limite et ainsi pouvoir considérer un déplacement des particules pendant leur durée de vie que les processus birth-death-move ont été introduits [LLG21].

Conformément à l'objectif de cette thèse, nous étudierons les processus birth-death-move afin de modéliser la dynamique spatio-temporelle d'un système de particules en mouvement au cours du temps, tandis que de nouvelles particules peuvent

apparaître dans la cellule et que certaines particules existantes peuvent disparaître. Dans ce modèle, les trajectoires peuvent être générées par n'importe quel modèle de diffusion de Markov continu, ce qui inclut la plupart des modèles de trajectoires individuelles précédemment considérés dans la littérature. De plus, ce modèle permet d'introduire des interactions entre les particules, de sorte que des situations de colocalisation (i.e. présence simultanée de plusieurs types de molécules en interaction au même endroit) peuvent ainsi être générées. L'intensité des naissances, qui régit le temps d'attente avant l'apparition de la prochaine particule, peut dépendre de la configuration actuelle des particules, de même que l'intensité des morts. Par conséquent, nous pouvons imposer que plus il y a de molécules dans la cellule, plus l'intensité de mort est élevée, ce qui induit des disparitions rapides. Certains effets spatiaux peuvent également être pris en compte afin de créer des régimes de mouvement distincts dans certaines régions de la cellule ou d'encourager certaines régions spatiales à concentrer davantage d'apparitions de particules, par exemple à proximité de certaines particules existantes.

Dans cette thèse, nous nous intéresserons aux applications du processus BDM en biologie cellulaire, mais ce type de dynamique est également observé dans d'autres domaines, notamment en épidémiologie [MH17] et en écologie [RS01, PG19]. Les processus BDM ont été introduits dans [LLG21] et analysés plus en détail dans [LLGM22], avec un accent particulier sur l'étude de leurs propriétés ergodiques. Pour être complet, plusieurs processus existants présentent des dynamiques pouvant être décrites comme des processus BDM : les systèmes de particules markoviens à sauts étudiés dans [Löc02b, Löc02a], les processus de branchement [AN12] et les modèles de population à structure spatiale [BM15].

Une première contribution de cette thèse a été d'ajouter au modèle existant la possibilité de marquer chaque particule avec une étiquette, permettant ainsi de distinguer, par exemple, différents types de protéines, ou différents types de régimes de mouvement suivis. Cette marque peut changer au fil du temps ; une particule peut, par exemple, transiter d'un régime brownien vers un mouvement dirigé. Ces changements de marque seront appelés, selon le contexte, *mutations* ou *transformations* (sans être une véritable mutation biologique). Ce modèle BDM marqué sera appelé processus birth-death-move avec mutations : BDMM.

La figure 2 présente les mêmes données que celles de la figure 1 avec l'ajout d'une

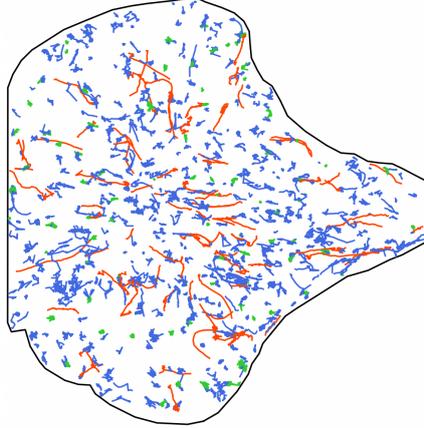


FIGURE 2 – Trajectoires obtenues après traitement de l’ensemble de la vidéo considérée dans la figure 1. Les couleurs indiquent la marque de chaque particule, à savoir le régime de mouvement suivi (bleu pour un mouvement brownien, rouge pour un mouvement super-diffusif et vert pour un mouvement confiné).

marque pour chaque particule, correspondant au type de mouvement. Cette marque est mise en évidence avec une couleur : bleu pour un mouvement brownien, rouge pour un mouvement super-diffusif et vert pour un mouvement confiné. Certaines particules peuvent changer de couleur, c’est-à-dire changer de type de mouvement au cours de leur durée de vie.

Dans cette figure, l’estimation des marques et des changements de marque a été effectuée par la méthode présentée dans [BVK19] et nous reviendrons sur cette problématique dans la dernière partie du manuscrit.

La dynamique d’un processus BDMM est expliquée rigoureusement dans le manuscrit. L’idée générale est la suivante. On note $\Lambda \subset \mathbb{R}^d$ l’espace de vie des particules considérées et \mathbb{M} l’ensemble fini des marques possibles. Un processus BDMM $(\mathbf{X}_t)_{t \geq 0}$ est un processus de Markov dont l’espace des états est défini ainsi :

$$E = \{\text{configurations de points de } \Lambda \times \mathbb{M}\} = \bigsqcup_{p=0}^{+\infty} E_p$$

où $E_p = \{\text{configurations de } p \text{ points de } \Lambda \times \mathbb{M}\}$. Au cours du temps, les particules se déplacent selon des mouvements continus entre deux “sauts”. Les “sauts” ont lieu à des instants aléatoires et correspondent à :

- une naissance : une nouvelle particule apparaît et le processus passe de l’espace

E_p à l'espace E_{p+1} ,

- une mort : une particule existante disparaît et le processus passe de l'espace E_p à l'espace E_{p-1} ,
- une mutation : la marque d'une particule existante change et le processus continue d'évoluer dans E_p .

Ainsi, pour entièrement décrire la dynamique d'un processus BDMM, trois ingrédients sont nécessaires :

- un processus de Markov multivarié, homogène et continu qui dirige le mouvement de toutes les particules entre deux instants de naissance, mort ou mutation ;
- trois fonctions, appelées intensités de naissance, de mort et de mutation, qui contrôlent les temps d'attente avant la prochaine naissance, la prochaine mort, la prochaine mutation ;
- trois noyaux de probabilité qui caractérisent, lorsqu'un saut a lieu, la loi de naissance de la particule qui apparaît, la loi de mort de la particule qui disparaît, et la loi de mutation de la particule qui change de marque.

Chacune de ces caractéristiques peut prendre une forme paramétrique, que l'on pourra chercher à estimer lors d'études statistiques.

1.3 État de l'art et contributions de la thèse

La première partie de ce manuscrit explique dans un premier temps, comment spécifier tous les ingrédients essentiels à la définition d'un modèle BDMM en donnant de nombreux exemples pertinents pour les applications en biologie cellulaire. Dans un deuxième temps, nous proposons un générateur du trafic de particules basé sur le processus BDMM.

L'outil de simulation est très flexible et permet de modéliser un système de particules en mouvement, éventuellement en interaction, chacune pouvant passer d'un régime de diffusion (par exemple, brownien) à un autre (par exemple, un mouvement dirigé), mais aussi de prendre en compte de manière précise l'apparition au cours du temps de nouvelles trajectoires et leurs disparitions, ces événements pouvant dépendre des régions dans la cellule mais aussi de la configuration spatiale courante de toutes les particules existantes. Notre simulateur permet

de générer des dynamiques plus réalistes que les procédures de simulation standards [SFS+97, GCK+18, AABA10, BKB09, Bou23, LCB+20], en tenant compte par exemple des situations de colocalisation souvent observées entre les protéines.

La deuxième partie du manuscrit est consacrée à l'inférence d'un processus BDMM. Dans des travaux antérieurs, Frédéric Lavancier et Ronan Le Guével ont décrit dans [LLG21] comment obtenir un estimateur non-paramétrique des fonctions d'intensité de naissance et d'intensité de mort (qui permettent de gérer le temps d'attente avant la prochaine naissance ou la prochaine mort). Dans le même état d'esprit, ils ont aussi, avec Emilien Manent dans [Man23], obtenu un estimateur non paramétrique des noyaux de naissance et de mort d'un processus birth-death-move.

Dans cette thèse, nous nous sommes concentrés sur l'aspect paramétrique de l'inférence du processus. Lorsqu'il n'y a ni déplacement ni mutation (i.e. lorsque le processus se résume à un processus spatial de naissance et de mort [Pre75]), l'estimation paramétrique par maximum de vraisemblance lorsque l'espace de vie des individus Λ est de dimension 1 a été examinée dans [MS94], avec une application à l'analyse du déplacement des dunes. La même méthode d'estimation a été reprise dans [Sad19] pour gérer des cas particuliers d'un processus spatial de naissance et de mort en dimension 2, en vue de modéliser les ouvertures et fermetures de magasins à Tokyo. Dans ces références avec des applications très différentes, aucune étude statistique théorique n'est fournie et, par définition, les particules ne se déplacent pas. Dans [Löc02a], la propriété de normalité asymptotique locale (LAN) du modèle markovien décrivant des systèmes de particules à sauts a été établie, et ceci dans un cadre très général.

La seconde contribution de cette thèse est d'avoir spécifié l'expression de la vraisemblance et démontré la propriété de normalité asymptotique locale (LAN) d'un processus BDMM sous certaines hypothèses, dont nous remarquerons qu'elles sont satisfaites pour les spécifications paramétriques standards du modèle. Cette propriété permettra d'obtenir, pour des approches paramétriques, des estimateurs asymptotiquement optimaux et de connaître de manière explicite leur loi asymptotique.

Les études statistiques, paramétriques ou non, nous permettent d'estimer chaque caractéristique du système sauf les instants de saut. Quant aux sauts qui corres-

pondent à des événements de naissance et de mort, il sont facilement identifiables en analysant le cardinal de la population. En revanche, l'estimation des instants de mutation demande une étude plus approfondie. La dernière partie de ce manuscrit se concentre sur cette étude, dans le cas particulier où l'ensemble \mathbb{M} des marques regroupe les différents régimes de déplacement dans lesquels peut se trouver chaque particule. Déterminer les instants de mutation revient donc, dans ce cas, à retrouver les instants où chaque particule change de régime de mouvement. L'objectif de cette dernière partie est donc d'estimer, à partir des trajectoires observées, le type de mouvement suivi par chaque particule et les différents paramètres sous-jacents, mais aussi de déterminer les instants de transition entre deux types de mouvement.

En général, la plupart des méthodes recourent au déplacement quadratique moyen (MSD : *Mean Square Displacement*) [GLGW13, QSE91] pour estimer le mouvement de diffusion (et ses paramètres) associé à une trajectoire complète. Des méthodes ont été élaborées pour détecter les instants de transition entre deux types de mouvement sur une même trajectoire ; certaines se basent sur des tests statistiques réalisés sur des fenêtres glissantes le long de la trajectoire [BVK19], d'autres sur la notion de chaîne de Markov à états cachés [MBP⁺15], et les plus récentes sur des techniques de "Deep Learning" [MGVGM⁺21, ASP⁺19].

La méthode que nous avons adoptée dans ce manuscrit repose sur des modèles de Markov à états cachés et un algorithme EM (Expectation-Maximization). Contrairement aux méthodes précédentes, notre approche permet à la fois de détecter les changements de type de mouvement mais aussi de déterminer les paramètres qui caractérisent ces mouvements, le tout en considérant un grand nombre de trajectoires. Pour finir, nous avons adapté notre approche pour gérer le cas où les particules suivent trois types de mouvement : un mouvement brownien, un mouvement brownien dirigé (super-diffusion) et un mouvement donné par un processus d'Ornstein-Uhlenbeck (sous-diffusion).

2 Organisation de la thèse

Cette thèse est organisée en trois parties, chacune étant subdivisée en deux chapitres avec l'ajout d'annexes qui s'y rapportent. Dans la partie I, nous présentons le modèle BDMM et un algorithme permettant de simuler des trajectoires de particules dans des cellules. Dans la partie II, nous étudions la propriété de normalité

asymptotique locale (LAN) du processus BDMM, avant de l'appliquer à l'estimation de certains paramètres du modèle. Enfin, dans la partie III, nous utilisons les modèles de Markov à états cachés et un algorithme EM pour estimer le mouvement suivi par un ensemble de particules. Le contenu de la thèse est présenté de manière synthétique ci-dessous.

Partie I : Un modèle génératif pour simuler la dynamique spatio-temporelle des biomolécules dans les cellules.

Cette partie correspond à l'article [BLSK23], publié dans le journal "Biological Imaging" en 2023. Pour cette raison, cette partie est rédigée en anglais.

Chapitre 1 : Présentation du processus birth-death-move avec mutations

Dans ce chapitre, nous étudions le processus birth-death-move avec changements de régimes (abrégé en BDMM) afin de générer des dynamiques conjointes de molécules dans les cellules.

Dans la première section, nous donnons la définition précise du processus stochastique BDMM, et nous énumérons, en nous appuyant sur de nombreux exemples, tous les ingrédients nécessaires pour décrire précisément le modèle. Une construction itérative est présentée, clarifiant le déroulement de la dynamique spatio-temporelle du processus. Un algorithme de simulation est détaillé en Annexe A et est disponible dans le répertoire Github <https://github.com/balsollier-lisa/BDM-generator-for-bioimaging>. Dans un deuxième temps, nous fournissons de nombreux exemples, que nous pensons être appropriés pour modéliser des dynamiques réelles de molécules.

Chapitre 2 : Application à l'étude de la dynamique de données réelles

En nous basant sur des ensembles de données réelles de microscopie de fluorescence, nous calibrons notre modèle pour mimer les dynamiques conjointes de deux types de protéines, Langerin et Rab-11, en interaction au niveau de la membrane plasmique. Notre approche nécessite une étude empirique préalable afin de choisir et de calibrer judicieusement les différents paramètres du modèle BDMM. Les protéines de type Langerin issues de ces données réelles sont celles représentées à droite

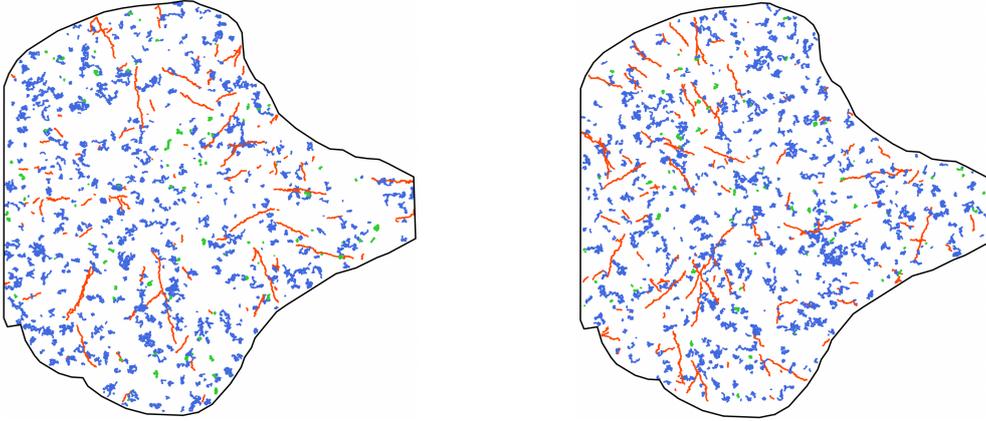


FIGURE 3 – Deux séquences simulées de trajectoires imitant la dynamique des protéines présentées dans la figure 1. Les couleurs indiquent la marque de chaque particule, qui est ici le régime de mouvement suivi (bleu pour un mouvement brownien, rouge pour un mouvement super-diffusif et vert pour un mouvement confiné).

de la figure 1. Après post-traitement, la séquence met en évidence une série de trajectoires, chacune suivant des types de mouvement distincts. Elles sont distribuées spatialement dans la cellule de manière spécifique, et se produisent à des intervalles variables tout au long de la séquence. De plus, l'apparition de nouvelles particules Langerin dépend de la position des protéines de type Rab-11 présentes. Cette description correspond bien à la dynamique d'un processus BDMM. Les trajectoires observées pour la protéine Langerin sont illustrées sur la figure 2, qui ajoute, par rapport à la figure 1, une marque pour chaque particule qui est ici son type de mouvement.

Par la suite, nous calculons, à partir de ce jeu de données réelles, un ensemble de descripteurs destinés à calibrer les paramètres du processus BDMM en vue de produire des simulations imitant la dynamique réelle. Ces descripteurs nous permettront aussi d'établir des caractéristiques de référence pour évaluer la pertinence des séquences simulées obtenues. Enfin, nous simulons plusieurs séquences et démontrons qu'elles possèdent des caractéristiques comparables à celles observées dans les séquences réelles. Deux simulations typiques de trajectoires mimant la dynamique réelle sont présentées sur la figure 3. Nous comparerons ces deux simulations aux données réelles de la figure 2.

Partie II : Estimation paramétrique et propriété LAN d'un processus BDMM.

Cette partie correspond à l'article [BL24], soumis pour publication en 2024. Pour cette raison, comme pour la partie précédente, elle est rédigée en anglais.

Chapitre 3 : Vraisemblance et propriété LAN d'un processus BDMM.

Dans ce chapitre, nous abordons l'inférence paramétrique des processus birth-death-move avec mutations. Après avoir défini de manière plus formelle que dans le chapitre 1 le processus BDMM, sa dynamique et ses différents ingrédients, nous nous intéressons à l'inférence paramétrique de ces caractéristiques étant donnée une seule réalisation du processus, sur un intervalle de temps fini. Dans une deuxième section, nous exprimons la vraisemblance d'un processus BDMM, ainsi que des garanties théoriques en prouvant la normalité asymptotique locale (LAN) du modèle. En notant $\mathcal{L}_t(\vartheta)$ la vraisemblance du processus observé sur l'intervalle $[0, t]$ et paramétré par $\vartheta \in \mathbb{R}^\ell$, cette propriété LAN exprime que sous certaines conditions de régularité, on a pour tout $\vartheta \in \mathbb{R}^\ell$, $t > 0$, $h \in \mathbb{R}^\ell$ et $N \in \mathbb{N}^*$:

$$\log \frac{\mathcal{L}_{tN}(\vartheta + \frac{h}{\sqrt{N}})}{\mathcal{L}_{tN}(\vartheta)} = h^T M^{N,\vartheta}(t) - \frac{1}{2} h^T \langle M^{N,\vartheta} \rangle(t) h + \text{Rem}_{N,\vartheta,h}(t),$$

où $\text{Rem}_{N,\vartheta,h}(t)$ converge vers 0 en $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probabilité lorsque N tend vers $+\infty$, $M^{N,\vartheta}$ est une martingale qui sera précisée dans le chapitre et $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ est la loi du processus sachant que sa position initiale est $\mathbf{x}_0 \in E$, ie. $\mathbf{X}_0 = \mathbf{x}_0$. Cette propriété implique l'optimalité et la normalité asymptotique de l'estimateur du maximum de vraisemblance, avec une matrice de covariance asymptotique explicite et estimable.

Chapitre 4 : Analyse du phénomène de colocalisation sur des données simulées puis sur des données réelles.

Dans ce chapitre, une étude par simulation est réalisée, montrant les performances de l'estimateur du maximum de vraisemblance de certains paramètres du modèle, ainsi que l'estimation de la région de confiance associée. Le cadre est proche de celui de l'ensemble de données réelles analysé par la suite, afin de fournir certaines garanties sur la fiabilité des conclusions qui s'y rapportent. Ce jeu de données est le même que celui présenté au chapitre 2, à savoir une séquence vidéo montrant des protéines Langerin et des protéines Rab-11 en interaction dans une cellule vi-

vante donnée. Comme détaillé dans le chapitre 2, leurs dynamiques sont cohérentes avec la réalisation du processus BDMM. En nous appuyant sur ce modèle et l'estimation de ses paramètres, nous quantifions le degré de colocalisation des protéines Langerin avec les protéines Rab-11. L'ensemble des codes Python utilisés pour obtenir les simulations et les résultats associés, ainsi que ceux associés aux données réelles sont téléchargeables à partir du le répertoire Github <https://github.com/balsollier-lisa/Parametric-estimation-of-the-BDMM-process>.

Partie III : Analyse du mouvement des trajectoires et détection des instants de changement de régime.

Cette partie est rédigée en français. Elle correspond à un travail réalisé durant la fin de ma thèse.

Pour mener à bien l'estimation paramétrique présentée dans la partie précédente, il est nécessaire de connaître les instants de saut du processus. Les temps de saut qui sont des instants de naissance (resp. de mort) sont faciles à identifier car il suffit d'observer à quels moments le nombre d'individus présents a augmenté (resp. diminué) de une unité. En revanche, les instants de mutation du processus sont beaucoup plus difficiles à détecter. Dans toute cette partie, nous nous plaçons dans le cas où la marque désigne le type de diffusion suivi par la particule. Nous supposons de plus, que lorsqu'une particule se déplace entre deux sauts, ce mouvement se fait indépendamment des autres particules, et que lors d'une mutation (qui est un changement de type de diffusion ici), celle-ci a lieu indépendamment des positions et des marques des autres particules. Le but de ce chapitre est d'estimer les instants de mutation dans ce cas particulier mais aussi de déterminer les caractéristiques du régime suivi par chaque particule entre ces instants.

Chapitre 5 : Modèles de Markov à états cachés et algorithme EM pour classifier des trajectoires

On suppose dans ce chapitre qu'un ensemble de trajectoires a été observé à des instants discrets. Le but est de classifier les différents morceaux de chacune des

trajectoires en fonction du régime de mouvement sous-jacent, et de déterminer les paramètres des régimes de mouvement (coefficients de diffusion, drift, ...).

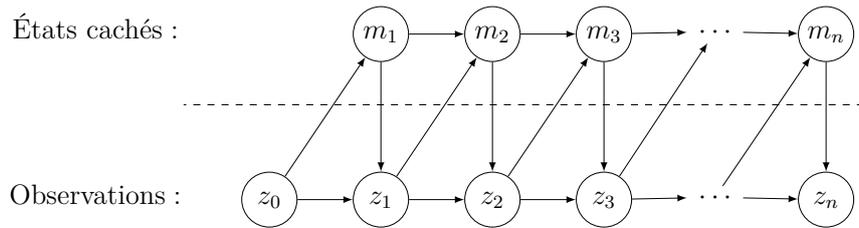
Pour chaque trajectoire, comportant n points d'un sous-ensemble $\Lambda \subset \mathbb{R}^d$, notée :

$$\mathbf{Z} = (z_0, z_1, \dots, z_n) \in \Lambda^{n+1},$$

on associe n états cachés, qu'il s'agit d'estimer :

$$\mathbf{M} = (m_1, \dots, m_n) \in \mathbb{M}^n,$$

où m_i représente le type de mouvement suivi par la particule entre le point z_{i-1} et z_i (e.g. brownien, confiné, super-diffusif). Ceci se traduit schématiquement par le graphe suivant :



Ce formalisme correspond à un modèle de Markov à états cachés. Nous recourons par la suite à un algorithme EM (Expectation-Maximization) pour estimer les paramètres optimaux de chacun des types de mouvement puis l'algorithme de Viterbi pour en déduire les états cachés. Dans le cas particulier où l'ensemble \mathbb{M} des différents types de mouvement possibles est $\mathbb{M} = \{\text{brownien, processus d'Ornstein-Uhlenbeck, brownien dirigé}\}$, nous montrons que les maxima nécessaires pour mettre en oeuvre l'algorithme EM sont explicites, ce qui rend d'autant plus efficace l'utilisation de cet algorithme.

Chapitre 6 : Applications sur des données simulées puis sur des données réelles.

Tout d'abord, nous nous intéressons à des données simulées, générées à l'aide de l'algorithme de simulation du modèle BDMM présenté dans la partie I, dont nous calibrons les caractéristiques pour mimer au mieux la dynamique des particules réelles que nous étudions dans la deuxième partie de ce chapitre. Cette dynamique présente les trois types de mouvement suivants : brownien, brownien dirigé et solution

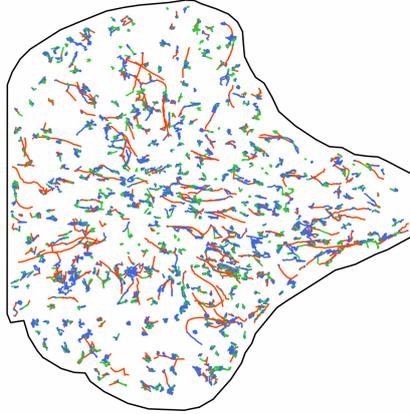


FIGURE 4 – Trajectoires de la figure 1 coloriées grâce aux résultats fournis par les algorithmes EM et Viterbi, développés dans la partie III. Les couleurs indiquent la marque de chaque particule, qui est ici le régime de mouvement suivi (bleu pour un mouvement brownien, rouge pour un mouvement brownien dirigé et vert pour un mouvement qui suit la solution d’un processus d’Ornstein-Uhlenbeck).

d’un processus d’Ornstein-Uhlenbeck. Nous présentons les résultats produits par l’algorithme EM, pour l’estimation des paramètres théoriques caractérisant ces trois mouvements dans divers scénarios. Par exemple, si toutes les trajectoires sont régies par les mêmes paramètres, l’estimation est de très bonne qualité. C’est également le cas si les paramètres propres à chaque trajectoire ne sont plus considérés comme des inconnues. En revanche, si les paramètres de chacun des trois mouvements sont propres à chaque trajectoire, ce qui revient au cas où on n’étudie qu’une seule trajectoire, les conclusions sont en deça des résultats espérés et d’autres méthodes plus dédiées sont plus efficaces [MBP⁺15].

Les résultats obtenus sur nos simulations s’avèrent précieux pour réaliser une étude similaire sur le jeu de données réelles déjà étudié dans les chapitres 2 et 4. Si on suppose que les paramètres de chacun des trois types de mouvement possibles sont communs à toutes les trajectoires, l’algorithme EM converge bien et permet d’estimer les états cachés (i.e. des types de mouvement suivis) grâce à l’algorithme de Viterbi. La figure 4 illustre le résultat d’estimation obtenu et procure ainsi une détection alternative à celle proposée par la méthode de [BVK19], présentée sur la figure 2.

PART I

**A generative model to simulate
the spatio-temporal dynamics of
biomolecules in cells.**

PRESENTATION OF THE BIRTH-DEATH-MOVE PROCESS WITH MUTATIONS

This chapter constitutes the first part of the article [\[BLSK23\]](#).

In this chapter, we give the precise definition of our stochastic process, and we in particular list all ingredients needed to fully specify the model. An iterative construction is presented in Section 1.2, clarifying how the dynamics proceeds, and an effective simulation algorithm is formally detailed in the appendix A. In Section 1.3, we provide numerous examples for the specifications of the model, that we think are relevant for many real biomolecules dynamics.

1.1 Heuristic and notations

In order to mimic the dynamics of biomolecules, we consider a multitype birth-death-move process with mutations, denoted by $(\mathbf{X}_t)_{t \geq 0}$. This process is a generalisation of birth-death-move processes, as introduced in [\[LLG21\]](#). In the following, to avoid misunderstanding we rather use the term “transformation” instead of “mutation”. This section describes the spatio-temporal dynamics of $(\mathbf{X}_t)_{t \geq 0}$ and introduces some notation.

At each time t , \mathbf{X}_t is a collection of particles located in a bounded set Λ of \mathbb{R}^d . Each particle is assigned a mark that represents a certain feature. We denote by \mathbb{M} the collection of possible marks. Through time, the particles move (possibly depending on their associated mark and in interaction with each other) and three sudden changes may occur, that we call “jumps” :

1. a “birth” : a new particle, assigned with a mark, may appear ;

2. a “death” : an existing particle may disappear ;
3. a “transformation” : the mark of an existing particle may change.

Example : In our biological application treated in Section 2.1 of Chapter 2, Λ represents a cell in dimension $d = 2$ or $d = 3$. We observe inside this cell two types of particles, associated to Langerin and Rab-11 proteins, and each of them moves according to three different possible regimes : Brownian, directed or confined motion. For this example, each particle is therefore marked out of 6 possibilities, whether it is associated to Langerin (L) or Rab-11 (R), and depending on its motion regime (1 to 3), so that $\mathbb{M} = \{(L, 1), (L, 2), (L, 3), (R, 1), (R, 2), (R, 3)\}$. Through time, each particle moves independently of the others according to its motion regime and eventually a new particle may appear, an existing one may disappear, and the motion regime of some particles may change.

We denote by $n(\mathbf{X}_t)$ the number of particles at time t . Each particle $x_i \in \mathbf{X}_t$, for $i = 1, \dots, n(\mathbf{X}_t)$, is decomposed as $x_i = (z_i, m_i)$ where $z_i \in \Lambda$ stands for its position while $m_i \in \mathbb{M}$ denotes its mark. Accordingly we have $\mathbf{X}_t \in (\Lambda \times \mathbb{M})^{n(\mathbf{X}_t)}$. (Strictly speaking the ordering of particles in \mathbf{X}_t does not matter, because any permutation of particles leads to the same collection of particles. We choose in this paper to bypass this nuance and use the same notation as if \mathbf{X}_t was a vector a particles, even it is actually a set of particles.) Since the number of particles changes over time, the stochastic process $(\mathbf{X}_t)_{t \geq 0}$ takes its values in the space

$$E = \bigcup_{n \in \mathbb{N}} (\Lambda \times \mathbb{M})^n.$$

To stress the fact that \mathbf{X}_t is not a simple value but encodes the positions and marks of a system of particles, we will say that this system at time t is in *configuration* \mathbf{X}_t .

To fully specify the dynamics of $(\mathbf{X}_t)_{t \geq 0}$, we need the following ingredients :

1. A system of equations (*Move*) that rules the way each particle of \mathbf{X}_t moves continuously between two jumps. We will typically consider a system of stochastic differential equations acting on the position of each particle, possibly depending on their associated mark and in interaction with the other particles ;
2. Three continuous bounded functions β , δ and τ from E to $[0, \infty)$, called birth, death and transformation intensity functions respectively, that govern the waiting times before a new birth, a new death and a new transformation. At each

time t , we may interpret $\beta(\mathbf{X}_t)dt$ as the probability that a birth occurs in the interval $[t, t + dt]$, given that the system of particles is in the configuration \mathbf{X}_t , and similarly for δ and τ .

3. Three transition probability functions that indicate how each jump occurs :
 - $p^\beta((z, m)|\mathbf{x})$: probability density function that the birth occurs at the position z with the mark m , given that there is a birth and that the system of particles is in configuration \mathbf{x} at the birth time ;
 - $p^\delta(x_i|\mathbf{x})$, for $x_i \in \mathbf{x}$: probability that the death concerns the particle x_i in \mathbf{x} , given that there is a death and that the system of particles is in configuration \mathbf{x} at the death time ;
 - $p^\tau((z_i, m_i), m|\mathbf{x})$, for $(z_i, m_i) \in \mathbf{x}$: probability that the particle $x_i = (z_i, m_i)$ in \mathbf{x} changes its mark and that this transformation leads to the new mark $m \neq m_i$, given that there is a transformation and that the system of particles is in configuration \mathbf{x} at the transformation time.

We provide in Section 1.3 some examples for the choice of these characteristics. Finally, we will denote by T_1, T_2, \dots the jump times of the process and we agree that $T_0 = 0$.

1.2 Algorithmic construction

Assume we are given the characteristics of the process $(\mathbf{X}_t)_{t \geq 0}$ as introduced in the previous section, that are the system of equations (*Move*), the intensity functions β, δ, τ , and the transition probability functions $p^\beta, p^\delta, p^\tau$. Then, starting from an initial configuration \mathbf{X}_0 at time $T_0 = 0$, we construct iteratively the process in the time interval $[0, T]$ as follows. Here we set $\alpha = \beta + \delta + \tau$ to be the total intensity of jumps.

1. Generate $n(\mathbf{X}_0)$ continuous trajectories as solutions of (*Move*) in the interval $[0, T - T_0]$, given the initial conditions \mathbf{X}_0 . Denote $(\mathbf{Y}_t)_{t \in [0, T - T_0]}$ these trajectories.
2. By flipping a coin, test whether the jump time T_1 occurs after T (this is with probability p) or before T (this is with probability $1 - p$), where

$$p = \exp \left(- \int_0^{T - T_0} \alpha(\mathbf{Y}_u) du \right).$$

- If $T_1 > T$, then $\mathbf{X}_t = \mathbf{Y}_{t-T_0}$ for all $t \in [T_0, T]$, which completes the simulation.
 - Otherwise, we continue by generating T_1 in $[T_0, T]$ and the associated jump as in the following.
3. Generate $T_1 - T_0$, given that $T_1 < T$, according to the probability distribution

$$\mathbb{P}(T_1 - T_0 < t | T_1 < T) = \frac{1}{1-p} \left(1 - \exp \left(- \int_0^t \alpha(\mathbf{Y}_u) du \right) \right), \quad 0 < t < T - T_0.$$

The process until the time T_1 is then given by the generated trajectories, i.e.

$$(\mathbf{X}_t)_{T_0 \leq t < T_1} = (\mathbf{Y}_{t-T_0})_{T_0 \leq t < T_1}.$$

4. Draw which kind of jump occurs at T_1 (we denote by $\mathbf{X}_{T_1^-}$ the configuration of the process just before the jump, which is $\mathbf{Y}_{T_1-T_0}$ by continuity of (\mathbf{Y}_t)):
- this is a birth with probability $\beta(\mathbf{X}_{T_1^-})/\alpha(\mathbf{X}_{T_1^-})$;
 - this is a death with probability $\delta(\mathbf{X}_{T_1^-})/\alpha(\mathbf{X}_{T_1^-})$;
 - this is a transformation with probability $\tau(\mathbf{X}_{T_1^-})/\alpha(\mathbf{X}_{T_1^-})$.
5. Generate the jump at $t = T_1$ to get \mathbf{X}_{T_1} as follows :
- if this is a birth, generate the new particle $x = (z, m)$ according to the probability density function $p^\beta((z, m) | \mathbf{X}_{T_1^-})$. Then set $\mathbf{X}_{T_1} = \mathbf{X}_{T_1^-} \cup (z, m)$;
 - if this is a death, draw which particle $x_i \in \mathbf{X}_{T_1^-}$ to delete according to the probability $p^\delta(x_i | \mathbf{X}_{T_1^-})$, for $i = 1, \dots, n(\mathbf{X}_{T_1^-})$. Then set $\mathbf{X}_{T_1} = \mathbf{X}_{T_1^-} \setminus x_i$;
 - if this is a transformation, draw which particle $x_i = (z_i, m_i) \in \mathbf{X}_{T_1^-}$ is transformed and generate its transformation according to $p^\tau((z_i, m_i), m | \mathbf{X}_{T_1^-})$, for $i = 1, \dots, n(\mathbf{X}_{T_1^-})$. Then set $\mathbf{X}_{T_1} = (\mathbf{X}_{T_1^-} \setminus (z_i, m_i)) \cup (z_i, m)$.
6. Back to step 1 with $T_0 \leftarrow T_1$ and $\mathbf{X}_0 \leftarrow \mathbf{X}_{T_1}$ in order to generate the new trajectories starting from \mathbf{X}_{T_1} and the next jump time T_2 , and so on.

In the first step of the above construction, the trajectories are generated up to the final time T . It is however very likely that the next jump occurs much before T so that it would be sufficient and computationally more efficient to generate these trajectories on a shorter time interval. We provide in Appendix A a formal algorithm of simulation of \mathbf{X}_t for $t \in [0, T]$, following the above construction and including the

latter idea. This algorithm has been implemented in Python and is available in our GitHub repository.

From a theoretical side, note that the specific exponential form of the probability distribution of the inter-jump waiting time in step 3 is necessary to imply the interpretation of β , δ and τ explained in the previous section. This exponential form also implies that $(\mathbf{X}_t)_{t \geq 0}$ is a Markov process, meaning that its future dynamics only depends on its present configuration. We refer to [LLG21] for more details about these theoretical aspects.

1.3 Exemplified specifications of the model

1.3.1 The inter-jumps motion

Recall that during an inter-jump period, the process (\mathbf{X}_t) has a constant cardinality $n = n(\mathbf{X}_t)$ and the marks of all its particles remain constant. We denote by (\mathbf{Y}_t) a system of n such particles $(z_{i,t}, m_i)$, for $i = 1, \dots, n$, where $z_{i,t} \in \Lambda$ represents the position of the i -th particle at time t and $m_i \in \mathbb{M}$ is its constant mark, that is

$$\mathbf{Y}_t = ((z_{1,t}, m_1), \dots, (z_{n,t}, m_n)).$$

In agreement with the construction of the previous section, the inter-jump trajectory of each particle of (\mathbf{X}_t) will coincide with the n trajectories of (\mathbf{Y}_t) during this period.

As a general example, we assume that (\mathbf{Y}_t) follows the following system of stochastic differential equation, starting at $t = 0$ at the configuration $\mathbf{y}_0 \in (\Lambda \times \mathbb{M})^n$,

$$(Move) : \begin{cases} dz_{i,t} = v_i(t, \mathbf{Y}_t)dt + \sigma_i(t, \mathbf{Y}_t)dB_{i,t}, & t \geq 0, \quad i = 1, \dots, n, \\ \mathbf{Y}_0 = \mathbf{y}_0, \end{cases}$$

where the drift functions v_i take their values in \mathbb{R}^d , the diffusion σ_i are non-negative functions, and $(B_{i,t})$, $i = 1, \dots, n$, are n independent standard Brownian motions in \mathbb{R}^d . Here, \mathbf{y}_0 , v_i and σ_i are free parameters to be chosen.

Some conditions on the drift and diffusion functions are necessary to ensure the existence and unicity of the solution of $(Move)$. This holds for instance if these functions are Lipschitz [Øks13], a condition met for the following examples. In addition, since each particle is supposed to evolve in the bounded set Λ of \mathbb{R}^d , we need in

practice to force the trajectories of (*Move*) to stay in Λ . This may be achieved by reflecting the trajectories at the boundary of Λ .

In its general form, (*Move*) allows the motion of each particle to depend on its mark, but also on the position and mark of the other particles (that are part of \mathbf{Y}_t). We detail several examples below, that may be realistic for biological applications.

Example 1 (Brownian motions). If $v_i(t, \mathbf{Y}_t) = 0$ and $\sigma_i(t, \mathbf{Y}_t) = \sigma$ (for $\sigma > 0$) is constant, then each particle follows a Brownian motion with the same diffusion coefficient σ , independently of the other particles.

Example 2 (spatially varying diffusion coefficients). If $v_i(t, \mathbf{Y}_t) = 0$ and $\sigma_i(t, \mathbf{Y}_t) = \sigma_{m_i}(z_{i,t})$, where σ_{m_i} is a positive function defined on Λ , then each particle follows an independent diffusive motion, where the diffusive coefficient depends on the associated mark and may vary in space. For instance, assume that $\Lambda = \Lambda_1 \cup \Lambda_2$ with $\Lambda_1 \cap \Lambda_2 = \emptyset$ and that for $m \in \mathbb{M}$,

$$\sigma_m(z) = \begin{cases} \sigma_{1,m} & \text{if } z \in \Lambda_1, \\ \sigma_{2,m} & \text{if } z \in \Lambda_2, \end{cases}$$

where $\sigma_{1,m} > 0$, $\sigma_{2,m} > 0$. Then each particle with mark m follows locally in Λ_1 a Brownian motion with diffusion coefficient $\sigma_{1,m}$ and locally in Λ_2 a Brownian motion with diffusion coefficient $\sigma_{2,m}$. Note that as such, σ_m is not Lipschitz and it needs to be smooth so as to fit the theoretical setting. This may be achieved by taking the convolution of σ_m by a bump function.

Example 3 (directed and confined motions). If $v_i(t, \mathbf{Y}_t) = v_{m_i}(z_{i,t})$ and $\sigma_i(t, \mathbf{Y}_t) = \sigma_{m_i}$, where v_{m_i} is defined on Λ and $\sigma_{m_i} > 0$, then each particle evolves independently of each other with a drift and a diffusion coefficient that depend on its mark. This example includes the directed motion considered in [BKV18] when $v_{m_i}(z) = v_{m_i}$ is a constant drift. It also includes the Ornstein-Uhlenbeck dynamics, also considered in [BKV18], when $v_{m_i}(z_{i,t}) = -\lambda_{m_i}(z_{i,t} - z_{i,0})$, where $\lambda_{m_i} > 0$ can be interpreted as a force of attraction towards the initial position $z_{i,0}$, leading to a confined trajectory.

Example 4 (interacting particles). In this example, we show how we can include interactions between the particles through a Langevin dynamics. To do so, we introduce, for $m, m' \in \mathbb{M}$, pairwise interaction functions $\Phi_{m,m'}$, as considered in statistical

physics : For $r > 0$, $\Phi_{m,m'}(r)$ represents the pairwise interaction between a particle with mark m and a particle with mark m' at a distance r apart. If $\Phi_{m,m'}(r) = 0$, there is no interaction, if $\Phi_{m,m'}(r) > 0$ there is inhibition between the two particles at distance r , and if $\Phi_{m,m'}(r) < 0$ there is attraction. Examples of inhibitive interaction functions can be found in [LLGM22]. The (overdamped) Langevin dynamics associated to these interactions reads as (*Move*) with $\sigma_i(t, \mathbf{Y}_t) = \sigma_{m_i}$, $\sigma_{m_i} > 0$, and

$$v_i(t, \mathbf{Y}_t) = - \sum_{j \neq i} \nabla \Phi_{m_i, m_j}(\|z_{i,t} - z_{j,t}\|),$$

where ∇ denotes the Gradient operator. Accordingly, each particle moves in a direction that tend to decrease the value of the pairwise interaction function with the other particles.

Example 5 (colocalized particles). Assume that some particles, say with mark m , are thought to be colocalized with particles having the mark \tilde{m} . This means that we expect the former to be localised nearby the latter and to follow approximately the same motion. Specifically, to let the particle i with mark m be colocalized with the particle j with mark \tilde{m} , we may simply define $z_{i,t} = z_{j,t} + \sigma_i B_{i,t}$, $t > 0$, where $B_{i,t}$ is a standard Brownian motion in \mathbb{R}^d representing the deviation of the trajectory i around the trajectory j , and $\sigma_i > 0$ quantifies the strength of this deviation. Here $z_{j,t}$ may be defined as in the previous examples, for instance as the typical trajectory of a particle with mark \tilde{m} .

1.3.2 The intensity functions

Recall that the intensity functions β , δ and τ rule the waiting times until the next birth, death and transformation, respectively. Heuristically, the probability that a birth occurs in the time interval $[t, t + dt]$ given that the particles are in configuration \mathbf{X}_t is $\beta(\mathbf{X}_t)dt$, and similarly for δ and τ . As a consequence these probabilities may evolve over time according to the configuration of particles, making for instance a death more likely to happen when there are many particles or a high concentration of them in some region, due to competition. We provide some natural examples below. For each example, any of β , δ or τ can be set similarly, even if we focus only on one of them.

Example 6 (constant intensities). The simplest situation is when the intensity

functions are constant, for instance $\beta(\mathbf{X}_t) = \beta$ with $\beta > 0$. Then births appear at a constant rate and we can expect that in average $\beta \times (s - t)$ new particles appear during the interval $[s, t]$.

Example 7 (intensities depending on the cardinality). If $\delta(\mathbf{X}_t) = \delta n(\mathbf{X}_t)$, with $\delta > 0$, then the more particles there are, the more deaths we observe. This is a natural situation when each particle is thought to have a constant death rate δ , so that the total death intensity for the system of particles at time t is just the sum of them, that is $\delta n(\mathbf{X}_t)$.

Example 8 (spatially varying intensities). Assume that the mark of a particle (say its motion regime) has more chance to change in some region of Λ than another, then the transformation intensity τ may reflect this dependency. Let for instance $\Lambda = \Lambda_1 \cup \Lambda_2$ with $\Lambda_1 \cap \Lambda_2 = \emptyset$ and define $\tau(\mathbf{X}_t) = \tau_1 n_1(\mathbf{X}_t) + \tau_2 n_2(\mathbf{X}_t)$ where $\tau_1 < \tau_2$ and $n_1(\mathbf{X}_t)$ (resp. $n_2(\mathbf{X}_t)$) denotes the number of particles in Λ_1 (resp. in Λ_2). Then for a given cardinality $n(\mathbf{X}_t)$, the more proportion of particles in Λ_2 , the more transformations happen. Note that in order to be rigorous, we should consider a continuous version of τ , which can be achieved by convolution with a bump function.

Example 9 (transformation due to colocalization). Assume that some m -particles (that are the particles with mark m) can be colocalized with some \tilde{m} -particles. Assume in addition that the particles are assigned a second mark that encodes their motion regime (e.g. diffuse, confined or directed). Eventually, during the dynamics of particles, a non-colocalized m -particle may become colocalized with a \tilde{m} -particle, meaning that it becomes D -close to a \tilde{m} -particle, where D is some prescribed colocalization distance. If so, we may expect that the motion regime of the m -particle becomes similar as the \tilde{m} -particle, so that a transformation must occur. Let $n_{D,m,\tilde{m}}(\mathbf{X}_t)$ be the number of D -close pairs of particles with marks m and \tilde{m} , whose motion regimes are different. Then we may define $\tau(\mathbf{X}_t) = \tau n_{D,m,\tilde{m}}(\mathbf{X}_t)$, for some $\tau > 0$, so that a transformation (of motion regime here) is very likely to occur when the aforementioned situation happen. Note that if τ is large, such transformation will quickly happen as soon as $n_{D,m,\tilde{m}}(\mathbf{X}_t) = 1$, and so $n_{D,m,\tilde{m}}(\mathbf{X}_t) > 1$ will be unlikely to be observed. Here again a smooth version of τ can be introduced by convolution to ensure its continuity.

1.3.3 The transition probability functions

We detail examples for the three possible transitions, in order below : births, deaths and transformations.

For the births, remember that $p^\beta((z, m)|\mathbf{x})$ denotes the probability density function (pdf) that a particle appears at the position z with the mark m , given that the system of particles are in configuration \mathbf{x} . To set this probability, two approaches are possible :

1. First drawing the mark m of the new particle with respect to some probability $p^\beta(m|\mathbf{x})$, then the position of the new particle given its mark according to some pdf $p^\beta(z|\mathbf{x}, m)$. This leads to the decomposition $p^\beta((z, m)|\mathbf{x}) = p^\beta(m|\mathbf{x})p^\beta(z|\mathbf{x}, m)$.
2. First generating the position of the new particle with respect to some pdf $p^\beta(z|\mathbf{x})$, then its mark m given the position with probability $p^\beta(m|\mathbf{x}, z)$. This leads to the decomposition $p^\beta((z, m)|\mathbf{x}) = p^\beta(z|\mathbf{x})p^\beta(m|\mathbf{x}, z)$.

Example 10 (uniform births). This is the simple example where the births do not depend on the environment, are uniform in space and the marks are drawn with respect to some prescribed probabilities p_m , where $\sum_{m \in \mathbb{M}} p_m = 1$. The two above approaches then coincide with $p^\beta(m|\mathbf{x}) = p^\beta(m|\mathbf{x}, z) = p_m$ and $p^\beta(z|\mathbf{x}, m) = p^\beta(z|\mathbf{x}) = 1/|\Lambda|$ for $z \in \Lambda$.

Example 11 (colocalized births). We adopt here the first approach above. We first draw the marks independently of the environment by setting $p^\beta(m|\mathbf{x}) = p_m$ with $\sum_{m \in \mathbb{M}} p_m = 1$, as in the previous example. Second, in order to generate the position of a new m -particle, thought to be colocalized with the \tilde{m} -particles, we may use a mixture of isotropic normal distribution, centered at each \tilde{m} -particle, with deviation $\sigma > 0$. Denoting by $\tilde{n}(\mathbf{x})$ the number of \tilde{m} -particles in \mathbf{x} and \tilde{z}_i their positions ($i = 1, \dots, \tilde{n}(\mathbf{x})$), this means that

$$p^\beta(z|\mathbf{x}, m) = \frac{1}{\tilde{n}(\mathbf{x})} \sum_{i=1}^{\tilde{n}(\mathbf{x})} \frac{1}{(\sigma\sqrt{2\pi})^d} \exp\left(-\frac{\|z - \tilde{z}_i\|^2}{2\sigma^2}\right). \quad (1.1)$$

Note that to be rigorous $p^\beta(z|\mathbf{x}, m)$ should be restricted to Λ with a proper normalisation, otherwise some particles might be generated outside Λ . We omit these details.

Example 12 (spatially dependent new marks). We may adopt the second approach by first generating a uniform position for the new particle, i.e. $p^\beta(z|\mathbf{x}) = 1/|\Lambda|$ for $z \in \Lambda$, and second by drawing the mark according to the generated position. Let for instance $\Lambda = \Lambda_1 \cup \Lambda_2$ with $\Lambda_1 \cap \Lambda_2 = \emptyset$ and set

$$p^\beta(m|\mathbf{x}, z) = \begin{cases} p_{1,m} & \text{if } z \in \Lambda_1, \\ p_{2,m} & \text{if } z \in \Lambda_2, \end{cases}$$

where $\sum_{m \in \mathbb{M}} p_{1,m} = \sum_{m \in \mathbb{M}} p_{2,m} = 1$. Then depending on the position, the distribution of the marks may be different.

We now focus on the death transition, namely the probability $p^\delta(x_i|\mathbf{x})$, for $x_i \in \mathbf{x}$, that the particle x_i in \mathbf{x} disappears when there is a death.

Example 13 (uniform deaths). The simplest example is when a death occurs uniformly over the existing particles, that is $p^\delta(x_i|\mathbf{x}) = 1/n(\mathbf{x})$ for $i = 1, \dots, n(\mathbf{x})$.

Example 14 (deaths due to competition). We may imagine that, due to competition, a particle is more likely to disappear if there are too many neighbours around it. Let $n_D(x_i)$ be the number of neighbouring particles around x_i within distance $D > 0$. To take into account the competition at distance $D > 0$, we may define $p^\delta(x_i|\mathbf{x}) = n_D(x_i) / \sum_{j=1}^{n(\mathbf{x})} n_D(x_j)$. Similarly, if relevant, we may count the number of neighbours of a certain mark only.

Finally, we focus on $p^\tau((z_i, m_i), m|\mathbf{x})$, for $(z_i, m_i) \in \mathbf{x}$, which is the probability that the particle $x_i = (z_i, m_i)$ in \mathbf{x} changes its mark from m_i to $m \neq m_i$, when a transformation happens. Similarly as for the birth transition probability, it is natural to decompose this probability as

$$p^\tau((z_i, m_i), m|\mathbf{x}) = p^\tau((z_i, m_i)|\mathbf{x})p^\tau(m|\mathbf{x}, (z_i, m_i)),$$

where $p^\tau((z_i, m_i)|\mathbf{x})$ represents the probability to choose the particle (z_i, m_i) in the configuration \mathbf{x} , in order to change its mark, and $p^\tau(m|\mathbf{x}, (z_i, m_i))$ is the probability to choose the new mark m given that the transformed particle is located at z_i with mark m_i .

Example 15 (transformations independent on the environment). A typical situation is when the particle to transform is drawn uniformly over the existing

particles, that is $p^\tau((z_i, m_i)|\mathbf{x}) = 1/n(\mathbf{x})$, and the transformation is carried out independently on the environment, according to a transition matrix with entries $p_{m,m'} \geq 0$, $m, m' \in \mathbb{M}$, representing the probability to be transformed from mark m to mark m' . Here, for all $m \in \mathbb{M}$, we assume $p_{m,m} = 0$ in order to ensure a genuine transformation, and of course $\sum_{m' \in \mathbb{M}} p_{m,m'} = 1$. With this formalism, we thus have $p^\tau(m|\mathbf{x}, (z_i, m_i)) = p_{m_i, m}$.

Example 16 (spatially dependent transformations). To make the previous example spatially dependent, introduce $p_m(z)$, a pdf in Λ representing the locations in Λ where a particle with mark m is more or less likely to be transformed. Then we may set

$$p^\tau((z_i, m_i)|\mathbf{x}) = \frac{n_i(\mathbf{x})}{n(\mathbf{x})} \frac{p_{m_i}(z_i)}{\sum_{j=1}^{n_i(\mathbf{x})} p_{m_i}(z_j^{(m_i)})},$$

where $n_i(\mathbf{x})$ denotes the number of particles with marks m_i in \mathbf{x} and $z_j^{(m_i)}$ for $j = 1, \dots, n_i(\mathbf{x})$, their positions. In this expression $n_i(\mathbf{x})/n(\mathbf{x})$ is a weight accounting for the prevalence of mark m_i in \mathbf{x} and the sum in the denominator is a normalisation so that the probabilities sum to 1. Note that if $p_m(z)$ is the uniform pdf on Λ , then we recover the uniform distribution $p^\tau((z_i, m_i)|\mathbf{x}) = 1/n(\mathbf{x})$. Furthermore, once the particle is chosen as above, we may apply a spatially dependent transformation as follows. Let $\Lambda = \Lambda_1 \cup \Lambda_2$ with $\Lambda_1 \cap \Lambda_2 = \emptyset$ and let two different transition matrices with respective entries $p_{m,m'}^{(1)}$ and $p_{m,m'}^{(2)}$, for $m, m' \in \mathbb{M}$. Then we may set

$$p^\tau(m|\mathbf{x}, (z_i, m_i)) = \begin{cases} p_{m_i, m}^{(1)} & \text{if } z_i \in \Lambda_1, \\ p_{m_i, m}^{(2)} & \text{if } z_i \in \Lambda_2. \end{cases}$$

Accordingly, the transformation does not follow the same distribution, whether the chosen particle to be transformed is located in Λ_1 or Λ_2 .

Example 17 (transformation due to colocalization). Assume that we are in the same situation as in Example 9 where m -particles can be colocalized to \tilde{m} -particles. We assume like in this example that a transformation occurs if $n_{D, m, \tilde{m}}(\mathbf{X}_t) \geq 1$, where $n_{D, m, \tilde{m}}(\mathbf{X}_t)$ denotes the number of D -close pairs of particles with marks m and \tilde{m} , whose motion regimes are different. Then, when a transformation happens, we may choose the m -particle to be transformed uniformly over those m -particles that are D -close to a \tilde{m} -particle with a different motion regime. Then the transformation makes the motion regime of the selected m -particle similar as the motion regime of

its closest \tilde{m} -particle.

1.4 Conclusion

In this chapter, we have presented the dynamics of the BDMM process, as well as an algorithm for simulation. This process is fully described by 3 main ingredients : 1) a stochastic equation that governs the motion of particles, 2) jump intensities that inform about the probability of a new particle appearing or a particle disappearing or changing regime, 3) transition kernels that control where new particles appear, select a particle to disappear or change regime to a new state. We then presented a number of examples to explain how these three ingredients can be set, depending on the target dynamics. In the next chapter, we present a methodological approach for calibrating the parameters involved in these three ingredients to faithfully represent the dynamics of Langerin and Rab-11 proteins at the surface of a cell.

APPLICATION TO THE STUDY OF THE DYNAMICS OF REAL DATA

This chapter constitutes the second part of the article [BLSK23]. Supplementary materials, including the Python code for simulation, the raw data and some further simulated sequences, are available in our online GitHub repository at <https://github.com/balsollier-lisa/BDM-generator-for-bioimaging>.

In this chapter, we demonstrate the potential of our model presented in the previous chapter for processing real data. In Section 2.1.1, we start by the inspection of a real dataset in order to judiciously calibrate the different parameters of the BDMM model to be in agreement with this example. In Section 2.1.2, we generate several simulated sequences and show that they exhibit comparable features as those observed in the real sequence.

2.1 Application to the joint dynamics of Langerin and Rab-11 proteins

2.1.1 Description of the dataset

The dataset we consider comes from the observation by a 3D multi-angle TIRF (total internal reflection fluorescence) microscopy technique of the intracellular trafficking of YFP Langerin and m-Cherry Rab-11 proteins in a RPE1 living cell [BGM⁺14], specifically projected along the z-axis onto the 2D plane close to the plasma membrane. This provides a 2D image sequence of 1199 frames, each lasting 140 ms and showing the simultaneous locations of the two types of proteins. The two images at the top of Figure 2.1 depict the first frame of the raw sequence for the Langerin fluorescent channel and the Rab-11 fluorescent channel, respectively, recorded simultaneously using a dual-view optical device. Note that the cell adheres on a

fibronectin micropattern, which constrains intracellular constituents such as cytoskeleton elements and gives a reproducible shape, explaining the “umbrella” shape of the cell. These raw sequences are post-processed following [PKB⁺08, PBB⁺14], then each bright spot is represented by a single point, and we apply the U-track algorithm [JLM⁺08] to estimate particle trajectories. The bottom images of Figure 2.1 show the resulting trajectories for the Langerin channel and the Rab-11 channel, respectively. These trajectories have been further analysed by the method developed in [BKV18] to classify them into three diffusion regimes : Brownian, directed and confined, which corresponds to the blue, red and green colors, respectively, in Figure 2.1. To be more specific in the analysis of all trajectories, we fit three parametric models to each of them, following [BKV18], depending on their regime :

1. For a Brownian regime (in blue) : a Brownian motion,
2. for a directed motion regime (in red) : a Brownian motion with constant drift,
3. for a confined motion regime (in green) : an Ornstein-Uhlenbeck process.

Each trajectory has its individual parameters, estimated by maximum likelihood [LS13]. Furthermore, some trajectories may change from one regime to another, which corresponds to a “transformation” in the BDMM model that will be specified in the next section.

Figure 2.2 summarises different features of the obtained trajectories for the Langerin sequence (the same characteristics have been analysed for the Rab-11 sequence, but are not detailed here). The histograms at the bottom display the duration of all trajectories (in frames), according to their regime. We can observe that the (blue) Brownian and (red) directed trajectories have quite a short lifetime in average, in comparison with the confined trajectories (in green). The top-right boxplots represent the distribution of the number of particles per frame, according to their regime : there is a majority of Brownian motions, followed by confined motions and a minority of directed motions. Finally, the top-left circular histogram aims at depicting the orientation of the drift vectors for the directed (red) trajectories. Specifically, for this plot, we have recorded the deviation of the drift angle (in degrees) with respect to the direction towards the center of the cell. For instance this deviation is 0° if the drift goes towards the center, and 180° if it goes in the opposite direction. It appears from this plot that most deviation angles are around 0° or 180° , meaning that the red trajectories mainly move in a radial direction going to (or starting from) the center of the cell.

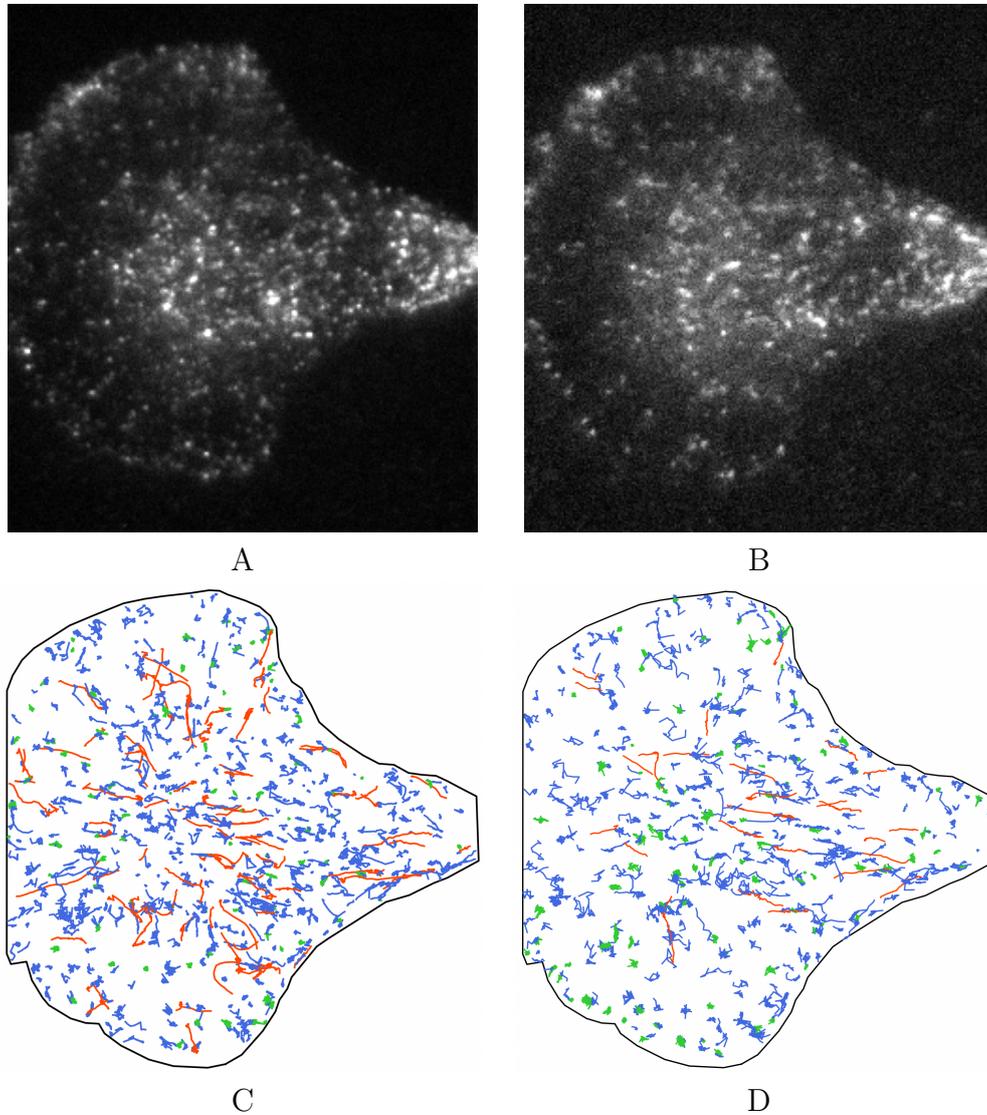


FIGURE 2.1 – A : first frame of the raw sequence showing in bright spots the location of Langerin proteins; B : same as A for the Rab-11 proteins; C : Set of all trajectories detected and tracked over the sequence of Langerin proteins colored by their estimated motion regime (Brownian in blue, directed motion in red and confined motion in green); D : same as C for the Rab-11 proteins

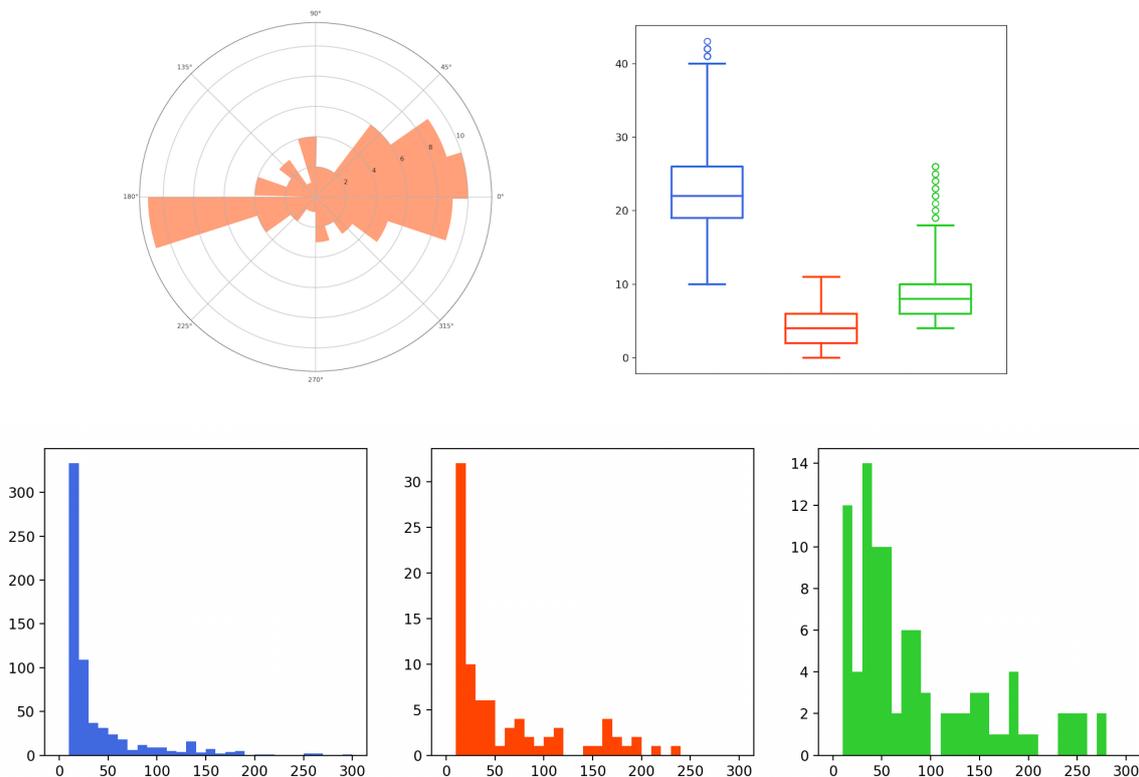


FIGURE 2.2 – Descriptors of the Langerin trajectories of the real-data sequence. Top-left : circular histogram of the deviation angle (from the direction towards the center of the cell) of the drifts of the directed trajectories. Top-right : boxplots of the number of trajectories per frame, according to their regime (blue : Brownian, red : directed motion, green : confined motion). Bottom : histograms of the lifetime (in frames) of each trajectory according to its regime (same color label)

The above descriptors will be helpful to calibrate the parameters of the BDM model in the next section and they will also serve as benchmarks to evaluate the quality of our simulations. However it is important to keep in mind that they come with some approximations and errors induced by imperfect tracking algorithms. In particular, no trajectory can last less than 10 frames in the data, which is a minimal length of detection for our tracking method. It is also clear in the bottom plots of Figure 2.1, that some directed trajectories appear wrongly in blue, which can be explained by the multiple testing procedure of [BKV18] that aims at minimizing the number of false positives (that are bad green or bad red trajectories) to the detriment of possibly too many false negatives (that are wrong blue trajectories).

TABLE 2.1 – Total number of births and deaths of trajectories observed in the realdataset sequence, according to the type of proteins and the motion regime.

		Brownian	Directed	Confined	Total	
Births	Langerin	603	78	66	747	1248
	Rab-11	393	24	84	501	
Deaths	Langerin	602	77	89	768	1282
	Rab-11	395	26	93	514	

TABLE 2.2 – Total number of regime transformations observed in the realdataset sequence of Langerin trajectories.

From \ To	Brownian	Directed	Confined
Brownian	0	0	9
Directed	0	0	1
Confined	5	0	0

Concerning the births and deaths of trajectories, we summarize in Table 2.1 their total numbers observed in the real-dataset, according to the type of proteins and motion regime. The number of regime transformations is in turn given in Table 2.2 for the Langerin proteins. For the Rab-11 proteins, only one switching from a Brownian motion to a confined motion was observed during the sequence.

To address in detail these jumps dynamics, we leverage the study carried out for the same dataset in [LLG21], where it has been concluded that for each type of proteins and motion regimes, the birth intensity is constant, like in Example 6 of Chapter 1, while the death intensity is proportional to the number of existing particles, like in Example 7. Given the small number of observed motion regime transformations, its intensity can also be considered as constant. Concerning the transition probability functions, the deaths occur uniformly over all existing particles, like in Example 13. As to the birth transition, there is no reason to choose another density than the uniform distribution over the cell for the Rab-11 proteins (Example 10). But due to colocalization (as observed for this dataset in [LPZK20]), the birth density for the Langerin positive structures can be approximated by a

mixture between a uniform distribution, for 93% of the Langerin births, and a co-localized density around the existing Rab-11 vesicles, like in Example 11, for 7% of the Langerin births. These proportions, along with the other parameters, have been estimated by maximum likelihood, as presented in Chapter 4. Note however that at this step, the goal is to provide a guideline to set the parameters of the BDM model in order to generate realistic realisations, as carried out in the next section. For this reason, any alternative estimation method or biological expertises to set the parameters could be appropriate.

2.1.2 Simulation of synthetic sequences

2.1.2.1 Model parameters setting

Based on the data analysis of the previous section, we are now in position to specify all characteristics of the BDMM process presented in Chapter 1, so as to mimic the joint dynamics of Langerin/Rab-11 proteins within a cell. To make the connection with the theoretical notation, the region of interest Λ represents the cell in dimension $d = 2$. Each particle in Λ will be marked by a label from the set $\mathbb{M} = \{(L, 1), (L, 2), (L, 3), (R, 1), (R, 2), (R, 3)\}$, where L stands for the Langerin proteins, R for the Rab-11 proteins, and the number 1, 2 or 3 indicates the motion regime of the particle : Brownian, directed or confined, respectively.

Concerning the motion of each trajectory, it follows the regime indicated by its mark and is in agreement with the observed trajectories from the real-dataset detailed in the previous chapter, see also Examples 1 and 3 :

- For a Brownian motion, we draw the diffusion coefficient according to the empirical distribution of the diffusion coefficients estimated from the Brownian motions of the real-dataset, for the same type of proteins (L or R) ;
- For a directed motion, we generate a Brownian motion with constant drift, with the same strategy for the choice of the diffusion coefficient, and where the drift vector is chosen as follows : it is by default oriented towards the center of the cell, this orientation being subjected to a deviation drawn from the empirical distribution depicted in the top-left circular histogram of Figure 2.2. In addition, its norm is drawn from the empirical distribution of the drift norms observed from the real-dataset. Here again, each set of parameters is distinct for the Langerin and Rab-11 proteins ;

- For a confined motion, we generate an Ornstein-Uhlenbeck process with diffusion coefficient 0.4 for all particles (which is the average from the real-dataset), and parameter $\lambda = 5.96$ for the Langerin proteins, and $\lambda = 7.41$ for the Rab-11 proteins.

In these values, the unit is pixels, and one pixel is $160 \times 160 \text{ nm}^2$ in our images.

Concerning the intensity functions, we set the birth intensity and the transformation intensity to constant values, as concluded from the real-data analysis. In agreement with Table 2.1, the total birth intensity can be estimated by $\beta(\mathbf{x}) = 1248/167.86 = 7.43$, whatever the configuration \mathbf{x} of particles is, because 1248 is the total number of observed births and 167.86 is the total time length of the sequence (in seconds). Similarly we set $\tau(\mathbf{x}) = 16/167.86 = 0.095$ since 16 transformations have been observed in the real sequence. For the death intensity, for each mark $m \in \mathbb{M}$, we let it proportional to the number of particles, that is $\delta_m n_m(\mathbf{x})$, where $n_m(\mathbf{x})$ is the number of particles with mark m in the configuration \mathbf{x} and δ_m has been estimated from the real-dataset as follows : $\delta_m = 0.17$ if $m = (L, 1)$, $\delta_m = 0.14$ if $m = (L, 2)$, $\delta_m = 0.07$ if $m = (L, 3)$, $\delta_m = 0.21$ if $m = (R, 1)$, $\delta_m = 0.25$ if $m = (R, 2)$, and $\delta_m = 0.08$ if $m = (R, 3)$. The total death intensity for the configuration \mathbf{x} of particles is then $\delta(\mathbf{x}) = \sum_{m \in \mathbb{M}} \delta_m n_m(\mathbf{x})$.

Finally we set the transition probability functions as follows. For the death transition, the probability to kill the particle $x_i = (z_i, m_i)$ in the configuration \mathbf{x} is set to

$$p^\delta((z_i, m_i)|\mathbf{x}) = \frac{\delta_{m_i}}{\delta(\mathbf{x})},$$

which means that we first draw the mark m with probability $\delta_m n_m(\mathbf{x})/\delta(\mathbf{x})$ and then the particle uniformly among all existing particles with mark m . For the transformation transition, we first select the type of proteins to transform with probability 15/16 for Langerin and 1/16 for Rab-11, in line with the transformations rates observed in the real-sequence, second we choose a particle uniformly among all existing particles of this type, and third, as in Example 15 of Chapter 1, we apply a regime transformation with respect to the following transition matrix (from the regime in rows to the regime in columns) :

$$\begin{array}{c} 1 \quad 2 \quad 3 \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \begin{pmatrix} 0 & 1/4 & 3/4 \\ 1/2 & 0 & 1/2 \\ 1 & 0 & 0 \end{pmatrix}. \end{array}$$

This matrix is in agreement with Table 2.2 concerning the Langerin proteins, where we have added some possible transitions from regime 1 to 2, and from regime 2 to 1, that appear to us likely to occur, even if they were not observed in the (quite rare) transformations in the real-sequence. The same transition matrix has been set for the Rab-11 proteins, since there is not enough observed transformations in the real-sequence (only one) to design a finer choice.

It remains to set the birth transition probability. First, we select which type of protein to create : following Table 2.1, it is a Langerin protein with probability $747/1248$ and a Rab-11 protein with probability $501/1248$. If the selected type is Rab-11, then it is generated uniformly in the cell with regime 1 with probability 0.784, 2 with probability 0.048 and 3 with probability 0.168, which corresponds to the relative proportion of births of each regime over all births for the real Rab-11 sequence. If the selected type is Langerin, then we flip a coin for colocalization with probability $p = 0.07$. If there is no colocalization, then the new Langerin protein is generated uniformly in the cell with regime 1 with probability 0.807, 2 with probability 0.104 and 3 with probability 0.088 (the observed relative proportions of births). If there is colocalization, then the new Langerin protein is generated around an existing Rab-11 protein according to the density (1.1) (page 35) in Example 11 of Chapter 1, where $\sigma = 1.1$. In this case, the regime of the new Langerin protein and its drift vector for a directed motion are similar as the those of its colocalized Rab-11 protein.

2.1.2.2 Analysis of resulting simulations

We have generated 100 sequences following the model of the previous section, during the same time length as the real-sequence of Section 2.1.1, that is 167.86 seconds for 1199 frames. Some descriptors concerning the generated Langerin trajectories coming from two simulated sequences are depicted in Figures 2.3 and 2.4, that are to be compared with the similar outputs of the real-data in Figures 2.1 and 2.2. The results for other simulated sequences can be seen in our GitHub repository. We have also summarized the mean number of births and deaths over the 100 simulated sequences in Table 2.3, to be compared with Table 2.1. Both graphical and quantitative results demonstrate that our model is able to create a joint dynamics with comparable features as those observed in the real-data sequence.

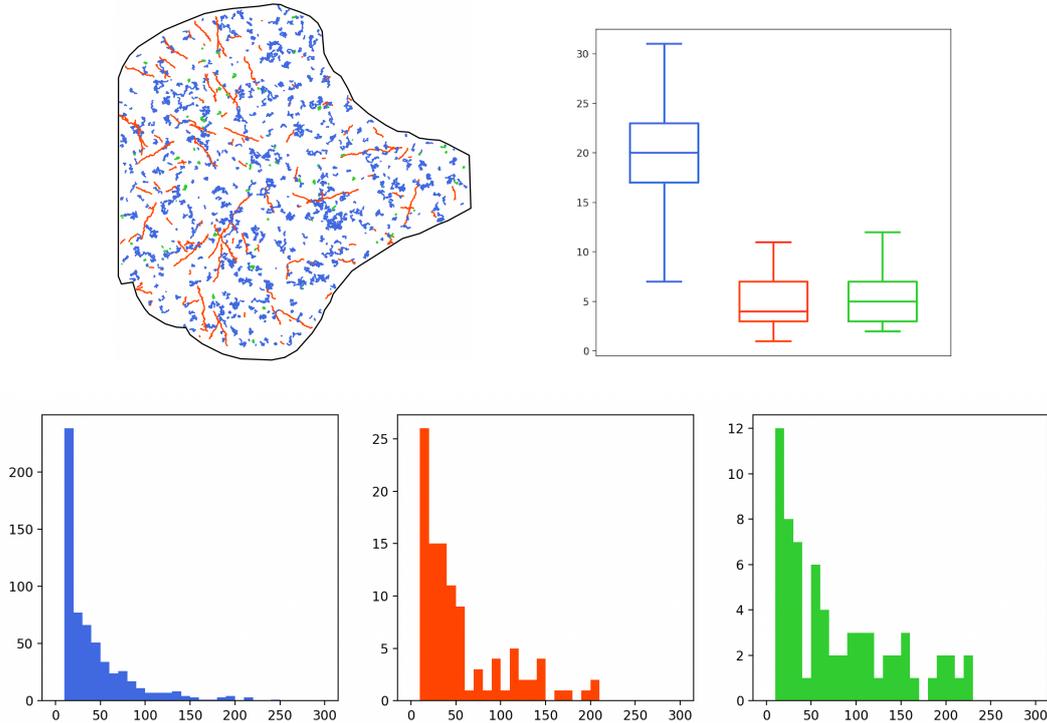


FIGURE 2.3 – Descriptors of the Langerin trajectories of a first simulated sequence. Top-left : set of trajectories, coloured according to their motion regime (blue : Brownian, red : directed, green : confined). Top-right : boxplots of the number of trajectories per frame, according to their regime (same color label). Bottom : histograms of the lifetime (in frames) of each trajectory according to its regime (same color label)

2.2 Conclusion and perspectives

We have leveraged an original stochastic model, namely a multitype birth-death-move process with transformations, in order to simulate biomolecules dynamics within a cell. This stochastic process not only models the individual trajectory of particles with any Markovian dynamics, but it is also able to generate the appearance (i.e., birth), disappearance (i.e., death) and regime switching (i.e., transformation) of each trajectory over time. Importantly, interactions between particles can be included, accounting for the possible colocalization phenomenon. The model is very flexible and is specified thanks to three sets of parameters : 1) a system describing the set of trajectories (typically a system of stochastic differential equations) ; 2) the intensity functions, ruling the waiting time before a new appearance, a disappearance or a switching ; 3) the transition probability functions, driving where a new

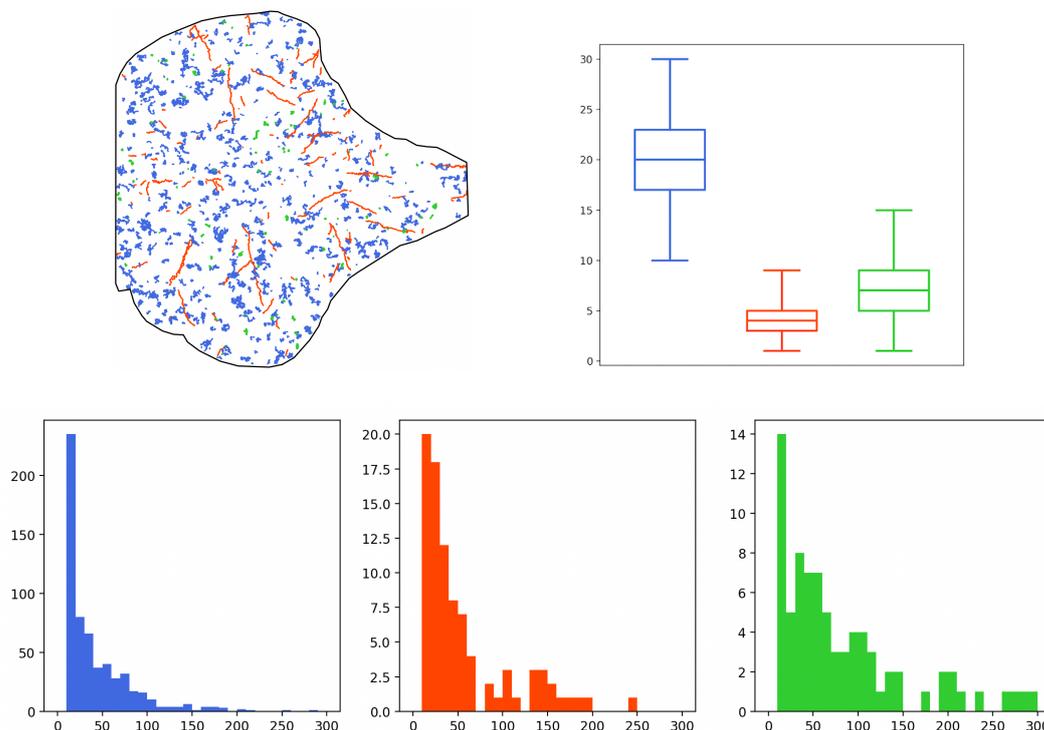


FIGURE 2.4 – Descriptors of the Langerin trajectories of a second simulated sequence, as in Figure 2.3

particle appears when there is a birth, which particle disappears when there is a death, and which particle switches its regime (and how) when there is a transformation. Numerous examples of these model specifications have been detailed. As an illustration, we demonstrated the relevance of this approach by generating the joint dynamics of Langerin/Rab-11 proteins within a cell, based on a preliminary data-based analysis in order to finely calibrate the model.

Since the model is very flexible, an important step is the choice of model characteristics and parameters. The calibration carried out for our illustration is specific to the application at hand, and of course another calibration must be carried out for another application. In our case, we used one observed sequence of Langerin/Rab-11 proteins. In order to improve the choice of parameters, a deeper empirical study based from several image sequences might help calibrating robustly the model. Once the parameters are fitted, the simulation of a sequence is quite fast : about one minute on a regular laptop for the generation of 2000 frames containing each about 70 trajectories in interaction. In general terms, the bottleneck is the simulation of

TABLE 2.3 – Mean total number of births and deaths of trajectories per sequence, over 100 simulated sequences.

		Brownian	Directed	Confined	Total	
Birth	Langerin	581	85	76	742	1239
	Rab-11	391	22	83	496	
Death	Langerin	584	86	77	747	1248
	Rab-11	397	23	80	500	

all trajectories between two jumps : if each particle moves independently of the others, this scales linearly with the number of particles and parallelization is easy to set up. When complicated interactions are introduced between particles, then the simulation of all trajectories scales badly with the number of particles. As a restriction, due to the Markovian framework ensuring the theoretical well-posedness of the model, anomalous trajectories [[ASP+19](#), [MGVGM+21](#), [KWFA18](#)] are not allowed in theory, though the algorithmic construction in Section 1.2 of Chapter1 does not rule out their introduction. However a rigorous understanding of the model in this setting remains challenging and constitutes an exciting perspective. In an effort to generate even more realistic image sequences, we may consider to blur the system of generated particles using the point spread function, and to add some noise and background, as carried out for instance in [[LCB+20](#)]. In relation, additional features could be computed from both the real-image sequence and the synthetic ones in order to strengthen the quality assessment of the generator.

Appendix of Part I

SIMULATION ALGORITHM

We provide in this appendix a formal algorithm to simulate a birth-death-move process with mutations (or transformations), following the construction of Section 1.2 of Chapter 1. Its implementation is available in our GitHub repository at <https://github.com/balsollier-lisa/BDM-generator-for-bioimaging> and in the second section of this appendix. It is a refinement of the algorithm introduced in [LLG21] for a birth-death-move process (without transformations). The idea is to generate the inter-jump move on a small time length Δ , then to test whether a jump has occurred during this period (this is with probability p in the following Algorithm 1). If so, we generate the jump time and the jump. If not, we continue the simulation of the inter-jump move on a further time length Δ , test whether a jump has occurred, and so on. The algorithm is valid whatever $\Delta > 0$ is, but an efficient choice is to set a small value for Δ . A default recommendation is to set Δ as the discretisation step used to simulate the trajectories in the *Algo.Move* algorithm (which is an input of our algorithm, see below).

We let as in Chapter 1 $\alpha = \beta + \delta + \tau$ and $T_0 = 0$. We denote in Algorithm 2 $\mathbf{X}_{T_j^-}$ the configuration just before the jump time T_j . In order to run Algorithms 1 and 2, we need the following inputs :

- $T > 0$: final time of simulation ;
- $\mathbf{X}_0 \in E$: initial configuration of particles ;
- $\Delta > 0$: small time length for piecewise simulation ;
- β, δ, τ : intensity functions of births, deaths and transformations ;
- $p^\beta, p^\delta, p^\tau$: transition probability for a birth, a death and a transformation ;
- *Algo.Move*(\mathbf{y}_0, Δ) : algorithm that returns, for $\mathbf{y}_0 \in E$ and $\Delta > 0$, $n(\mathbf{y}_0)$ (discretised) trajectories on $[0, \Delta]$ following the system of SDEs (*Move*) with initial configuration \mathbf{y}_0 .

Algorithm 1: Simulation on the time interval $[0, T]$

```

set  $t = 0$  and  $j = 0$  ;
while  $t < T$  do
    set  $k = 1$  ;
    while  $k \neq 0$  do
        set  $\Delta_j = \min(\Delta, T - T_j - (k - 1)\Delta)$  ;
        generate  $(\mathbf{Y}_s)_{0 \leq s \leq \Delta_j}$  by Algo.Move $(\mathbf{X}_{T_j + (k-1)\Delta}, \Delta_j)$  ;
        set  $p = \exp\left(-\int_0^{\Delta_j} \alpha(\mathbf{Y}_u) du\right)$  ;
        generate  $U \sim U([0, 1])$ ;
        if  $U \leq p$  then
            set  $\mathbf{X}_s = \mathbf{Y}_{s - T_j - (k-1)\Delta}$  for
             $s \in [T_j + (k - 1)\Delta, T_j + (k - 1)\Delta + \Delta_j]$  ;
            if  $\Delta_j = T - T_j - (k - 1)\Delta$  then
                 $k = 0$  ;
                 $t \leftarrow T$  ;
            else
                 $k \leftarrow k + 1$  ;
            end
        else
            generate the waiting time  $\tau$  according to the distribution

            
$$\mathbb{P}(\tau < s | (k - 1)\Delta < \tau < (k - 1)\Delta + \Delta_j)$$


$$= \frac{1 - \exp\left(-\int_0^{s - (k-1)\Delta} \alpha(\mathbf{Y}_u) du\right)}{(1 - p)}, \quad \forall s \in [(k-1)\Delta, (k-1)\Delta + \Delta_j];$$


            set  $T_{j+1} = T_j + \tau$  and  $\mathbf{X}_s = \mathbf{Y}_{s - T_j - (k-1)\Delta}$  for
             $s \in [T_j + (k - 1)\Delta, T_{j+1}[$  ;
            generate  $\mathbf{X}_{T_{j+1}}$  by Algorithm 2 given  $\mathbf{X}_{T_{j+1}^-} = \mathbf{Y}_{T_{j+1} - T_j - (k-1)\Delta}$  ;
             $k = 0$  ;
             $t \leftarrow T_{j+1}$  ;
             $j \leftarrow j + 1$  ;
        end
    end
end

```

Algorithm 2: Simulation of the jump at $t = T_{j+1}$ given $\mathbf{X}_{T_{j+1}^-}$

generate $U \sim U([0, 1])$;
if $U \leq \beta(\mathbf{X}_{T_{j+1}^-})/\alpha(\mathbf{X}_{T_{j+1}^-})$ **then**
 generate a new particle (z, m) according to $p^\beta((z, m)|\mathbf{X}_{T_{j+1}^-})$;
 set $\mathbf{X}_{T_{j+1}} = \mathbf{X}_{T_{j+1}^-} \cup (z, m)$;
else
 if $U \leq (\beta + \delta)(\mathbf{X}_{T_{j+1}^-})/\alpha(\mathbf{X}_{T_{j+1}^-})$ **then**
 sample $x_i \in \mathbf{X}_{T_{j+1}^-}$ according to $p^\delta(x_i|\mathbf{X}_{T_{j+1}^-})$;
 set $\mathbf{X}_{T_{j+1}} = \mathbf{X}_{T_{j+1}^-} \setminus (z, m)$;
 else
 sample $x_i = (z_i, m_i) \in \mathbf{X}_{T_{j+1}^-}$ and $m \in \mathbb{M}$ according to
 $p^\tau((z_i, m_i), m|\mathbf{X}_{T_{j+1}^-})$;
 set $\mathbf{X}_{T_{j+1}} = (\mathbf{X}_{T_{j+1}^-} \setminus (z_i, m_i)) \cup (z_i, m)$;
 end
end

PART II

**Parametric estimation and LAN
property of a BDMM process.**

LIKELIHOOD AND LAN PROPERTY OF A BDMM PROCESS.

This chapter constitutes the first part of the submitted article [BL24]. We start by presenting in a formal way the definition of the BDMM process. For this reason, some considerations may be redundant with Chapter 1, although the setting is slightly different. The main results are stated in Section 3.2 : we derive the likelihood expression of the model and we prove its LAN property under mild assumptions, that we prove to be satisfied for the examples presented in Section 3.1.3, under mild hypotheses. Appendices C and D contain the proofs and technical lemmas.

3.1 Definition of the BDMM process

3.1.1 State space

We denote by $(\mathbf{X}_t)_{t \geq 0}$ the BDMM process. At each time $t \geq 0$, \mathbf{X}_t describes the configuration of a set of marked particles. Each particle $x \in \mathbf{X}_t$ reads $x = (z, m)$ where $z \in \Lambda \subset \mathbb{R}^{p+q}$ encodes the spatial location of the particle in \mathbb{R}^p , $p \geq 1$, along with its possible continuous mark in \mathbb{R}^q , $q \geq 0$, and where $m \in \mathbb{M}$ is a discrete mark (or *label*), \mathbb{M} being the finite set of possible labels. In the following, we set $d = p + q$, so that $\Lambda \subset \mathbb{R}^d$.

At each time t , \mathbf{X}_t represents the set of all alive particles, where the ordering does not matter and the cardinality may change over time. For this reason, for any $n \geq 1$, we introduce the natural projection

$$\Pi_n : (x_1, \dots, x_n) \in (\Lambda \times \mathbb{M})^n \mapsto \{x_1, \dots, x_n\} \quad (3.1)$$

that identifies two elements (x_1, \dots, x_n) and (y_1, \dots, y_n) of $(\Lambda \times \mathbb{M})^n$ if there exists

a permutation π of $\{1, \dots, n\}$ such that $x_i = y_{\pi(i)}$ for any $1 \leq i \leq n$.

We set for any $n \geq 1$, $E_n = \Pi_n((\Lambda \times \mathbb{M})^n)$. The state space of \mathbf{X}_t for $t \geq 0$ is then

$$E = \bigsqcup_{n \geq 0} E_n,$$

where $E_0 = \{\emptyset\}$ consists of the empty configuration.

The cardinality of a configuration $\mathbf{x} \in E$ will be denoted by $n(\mathbf{x})$, so that $\mathbf{x} \in E_{n(\mathbf{x})}$, and we write

$$\mathbf{x} = \{x_1, \dots, x_{n(\mathbf{x})}\} = \{(z_1, m_1), \dots, (z_{n(\mathbf{x})}, m_{n(\mathbf{x})})\},$$

where $z_i \in \Lambda$ and $m_i \in \mathbb{M}$. In the following we will sometimes assimilate $(\Lambda \times \mathbb{M})^n$ with $\Lambda^n \times \mathbb{M}^n$ for convenience, so that we may also write $\mathbf{x} = \{(\mathbf{z}, \mathbf{m})\}$ where $\mathbf{z} = (z_1, \dots, z_{n(\mathbf{x})})$ and $\mathbf{m} = (m_1, \dots, m_{n(\mathbf{x})})$. Similarly, we will write the configuration at time t of the BDMM process in either way

$$\mathbf{X}_t = \{x_{1,t}, \dots, x_{n,t}\} = \{(z_{1,t}, m_{1,t}), \dots, (z_{n,t}, m_{n,t})\} \quad \text{or} \quad \mathbf{X}_t = \{(\mathbf{Z}_t, \mathbf{m}_t)\},$$

where $n = n(\mathbf{X}_t)$. Moreover, for $\mathbf{x} \in E_n$ and $(z, m) \in \Lambda \times \mathbb{M}$, we write $\mathbf{x} \cup (z, m)$ for $\{x_1, \dots, x_n, (z, m)\} \in E_{n+1}$, and for $n \geq 1$ and $1 \leq i \leq n$, we write $\mathbf{x} \setminus x_i$ for $\{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\} \in E_{n-1}$.

We equip E with the Borel σ -algebra \mathcal{E} and we endow E with the distance d_1 defined for \mathbf{x} and \mathbf{y} in E such that $n(\mathbf{x}) \leq n(\mathbf{y})$ by

$$d_1(\mathbf{x}, \mathbf{y}) = \frac{1}{n(\mathbf{y})} \left(\min_{\pi \in S_{n(\mathbf{y})}} \sum_{i=1}^{n(\mathbf{x})} (\|x_i - y_{\pi(i)}\| \wedge 1) + (n(\mathbf{y}) - n(\mathbf{x})) \right),$$

with $d_1(\mathbf{x}, \emptyset) = 1$ and where S_n denotes the set of permutations of $\{1, \dots, n\}$. This distance makes in particular the function $\mathbf{x} \mapsto n(\mathbf{x})$ continuous on E , and further implies some nice topological properties for (E, d_1) , see [SX08] and [LLGM22] for details.

3.1.2 Dynamics of a BDMM process

The dynamics of a BDMM process alternates continuous motions and jumps. Inter-jumps motions concern the location and/or the continuous mark of all alive

particles, that move in Λ . Jumps are of three types : either a birth (when a new particle appears), or a death (when an existing particle disappears), or a mutation (when an existing particle changes its label). Accordingly, the dynamics is based on the three following elements :

1. A continuous homogeneous Markov process $(\mathbf{Y}_t)_{t \geq 0}$ on E that drives the motion of all particles of $(\mathbf{X}_t)_{t \geq 0}$ between two jumps. This process will not affect the label of the particles, neither their cardinality. Given an initial condition $\mathbf{Y}_0 = \mathbf{x}$, $(\mathbf{Y}_t)_{t \geq 0}$ will be defined as the solution of a stochastic differential equation, as described and exemplified in Section 3.1.3.3.
2. Three continuous non-negative functions β , δ and τ on E , that refers respectively to the birth, death and mutation intensities. These functions govern the waiting times between the jumps of $(\mathbf{X}_t)_{t \geq 0}$. Heuristically, the probability that a birth occurs in the interval $[t, t + dt]$ given that the particles are in the configuration \mathbf{X}_t à time t is $\beta(\mathbf{X}_t)dt$, and similarly for δ and τ . We will denote $\alpha = \beta + \delta + \tau$ the total jump intensity of the process, that we assume to be bounded.
3. Three transition kernels K_β , K_δ , and K_τ from $E \times \mathcal{E}$ to $[0, 1]$, that specify how a jump occurs when it happens. For example, given that a birth occurs in a configuration \mathbf{x} , $K_\beta(\mathbf{x}, A)$ is the probability that the new configuration $\mathbf{x} \cup (z, m)$ belongs to $A \in \mathcal{E}$, where (z, m) denotes the new particle, and similarly for K_δ and K_τ .

Details and examples about these three basic characteristics of the process are provided in the next section.

Let us specifically describe the algorithmic definition of the process. This iterative construction follows [LLG21] and [LLGM22], and may serve as a simulation procedure, see [LLG21] for details. Let $(\mathbf{Y}_t^{(i)})_{t \geq 0}, i \geq 0$, be a sequence of processes on E , identically distributed as $(\mathbf{Y}_t)_{t \geq 0}$. Starting from the initial configuration \mathbf{X}_0 at time $T_0 = 0$, we iteratively build the process as follows.

- i) Given \mathbf{X}_0 , generate the $n(\mathbf{X}_0)$ continuous trajectories of $(\mathbf{Y}_t^{(0)})_{t \geq 0}$.
- ii) Given \mathbf{X}_0 and $(\mathbf{Y}_t^{(0)})_{t \geq 0}$, generate the first inter-jump time $T_1 - T_0$ according to the cumulative distribution function

$$F_1(t) = 1 - \exp\left(-\int_0^t \alpha(\mathbf{Y}_u^{(0)}) du\right), \quad t > 0. \quad (3.2)$$

The process until time T_1 is given by the generated trajectories, i.e.

$$(\mathbf{X}_t)_{T_0 \leq t < T_1} = (\mathbf{Y}_{t-T_0}^0)_{T_0 \leq t < T_1}.$$

iii) Given T_1 and $\mathbf{X}_{T_1^-} = \mathbf{Y}_{T_1-T_0}^0$, generate the first jump :

— it is a birth with probability $\beta(\mathbf{X}_{T_1^-})/\alpha(\mathbf{X}_{T_1^-})$, in which case

$$\mathbf{X}_{T_1} \sim K_\beta(\mathbf{X}_{T_1^-}, .);$$

— it is a death with probability $\delta(\mathbf{X}_{T_1^-})/\alpha(\mathbf{X}_{T_1^-})$, in which case

$$\mathbf{X}_{T_1} \sim K_\delta(\mathbf{X}_{T_1^-}, .);$$

— it is a mutation with probability $\tau(\mathbf{X}_{T_1^-})/\alpha(\mathbf{X}_{T_1^-})$, in which case

$$\mathbf{X}_{T_1} \sim K_\tau(\mathbf{X}_{T_1^-}, .).$$

iv) Return to step i) with $T_0 \leftarrow T_1$ and $\mathbf{X}_0 \leftarrow \mathbf{X}_{T_1}$ to generate the new trajectories $(\mathbf{Y}_t^{(1)})_{t \geq 0}$ starting from \mathbf{X}_{T_1} , the next jumping time T_2 , and so on.

The sequence of jumping times of the BDMM process $(\mathbf{X}_t)_{t \geq 0}$ is $(T_n)_{n \geq 1}$. For $t > 0$, we denote by N_t the number of jumps on $[0, t]$, i.e.

$$N_t = \sum_{i \geq 1} \mathbb{1}_{T_i \leq t}.$$

Similarly, we denote by N_t^β , N_t^δ and N_t^τ the number of births, deaths and mutations, respectively, before t . By our assumption that the total jump intensity α is bounded, we have that $N_t < \infty$ for any $t \geq 0$, i.e. there is no explosion of the process.

The BDMM process, as defined by the above iterative construction, is a particular case of a jump-move process, as studied in [LLGM22]. We deduce that it is a homogeneous Markov process with respect to its natural filtration $(\mathcal{F}_t)_{t \geq 0}$. We refer to the latter reference for the expression of the infinitesimal generator of the process and further probabilistic properties. In the following, we will denote by $\mathbb{P}_\mathbf{x}$ and $\mathbb{E}_\mathbf{x}$ all probabilities and expectations given that $\mathbf{X}_0 = \mathbf{x}$.

3.1.3 Elements of the model and examples

In this section, we describe more precisely each characteristics of the process, that are the intensity functions, the transition kernels, and the move process $(\mathbf{Y}_t)_{t \geq 0}$. We also provide various examples.

3.1.3.1 Intensity functions

The birth, death and mutation intensity functions are such that, given $\mathbf{X}_t = \mathbf{x}$, a birth (resp. a death and a mutation) occurs in $(t, t+h]$ with probability $\beta(\mathbf{x})h + o(h)$ (resp. $\delta(\mathbf{x})h + o(h)$ and $\tau(\mathbf{x})h + o(h)$) as $h \rightarrow 0$. These intensities can also be seen in the following way : $\beta(\mathbf{X}_{t-})$ (resp. $\delta(\mathbf{X}_{t-})$ and $\tau(\mathbf{X}_{t-})$) is the intensity of the associated counting process N_t^β (resp. N_t^δ and N_t^τ). These interpretations are consequences of the specific form (3.2) of the waiting time before the next jump, see [LLG21] for details.

To set some examples, let us denote by γ any of β , δ or τ . The most simple situation is when the intensity function is a constant rate, that is $\gamma(\mathbf{x}) = \gamma$ for any $\mathbf{x} \in E$, where $\gamma > 0$. Then there is in average $\gamma\Delta$ new events that occur in any time interval of size Δ , whatever the configuration of \mathbf{X}_t is. Another typical setting is when the intensity is proportional to the cardinality, that is $\gamma(\mathbf{x}) = \gamma n(\mathbf{x})$ for $\mathbf{x} \in E$ and $\gamma > 0$. This reflects the situation where each particle has its own rate $\gamma > 0$ and the particles do not interact, so that the total rate over all particles is $\gamma n(\mathbf{x})$. These two examples are observed in the biological applications studied in [LLG21] and [BLSK23]. More complicated examples of intensity functions are also considered in [LLG21], where $\gamma(\mathbf{x})$ depends on the underlying Voronoï tessellation induced by \mathbf{x} .

3.1.3.2 Transition kernels

Denoting γ for any of β , δ or τ , we assume that for any \mathbf{x} the transition kernel $K_\gamma(\mathbf{x}, \cdot)$ admits a density $k_\gamma(\mathbf{x}, \mathbf{y})$, $\mathbf{y} \in E$, with respect to a measure $\nu_\gamma(\mathbf{x}, \cdot)$, that is for all $\mathbf{x} \in E$ and $\mathbf{y} \in E$,

$$K_\gamma(\mathbf{x}, d\mathbf{y}) = k_\gamma(\mathbf{x}, \mathbf{y})\nu_\gamma(\mathbf{x}, d\mathbf{y}).$$

We now specify k_γ and ν_γ in each case, whether the transition concerns a birth, a death or a mutation.

For a birth transition, we assume that for any $F \in \mathcal{E}$, $\nu_\beta(\mathbf{x}, F) = 0$ except if there exist $A \subset \Lambda$ and $\mathcal{I} \subset \mathbb{M}$ such that

$$F = \mathbf{x} \cup (A \times \mathcal{I}) := \{\mathbf{y} \in E, \exists (z, m) \in A \times \mathcal{I}, \mathbf{y} = \mathbf{x} \cup (z, m)\}.$$

In this case we set $\nu_\beta(\mathbf{x}, \mathbf{x} \cup (A \times \mathcal{I})) = |A| \times |\mathcal{I}|$ to be the Lebesgue measure on $\Lambda \times \mathbb{M}$. This choice of ν_β ensures that the birth transition produces the addition of exactly one particle to the configuration \mathbf{x} , as expected for a birth. For the density $k_\beta(\mathbf{x}, \mathbf{y})$, we assume that

$$k_\beta(\mathbf{x}, \mathbf{y}) = \begin{cases} p_m k_\beta^m(\mathbf{x}, z) & \text{if } \exists (z, m) \in \Lambda \times \mathbb{M}, \mathbf{y} = \mathbf{x} \cup (z, m), \\ 0 & \text{otherwise,} \end{cases} \quad (3.3)$$

where $p_m \in [0, 1]$ is the probability that the new particle has the label $m \in \mathbb{M}$, with $\sum_{m \in \mathbb{M}} p_m = 1$, and $k_\beta^m(\mathbf{x}, \cdot)$ is the density for the location of the new particle in Λ given that its label is m .

A simple example of birth transition is the uniform law for the label and the location, which corresponds to $p_m = 1/|\mathbb{M}|$ for all $m \in \mathbb{M}$ and $k_\beta^m(\mathbf{x}, z) = 1/|\Lambda|$, for all $z \in \Lambda$ and $m \in \mathbb{M}$. More sophisticated examples are provided below.

Example 1 (Birth kernel as a mixture of normal laws). In this example, a new particle is more likely to appear close to existing particles. The birth density is a mixture of isotropic normal distributions, centred at each existing particle, with deviation $\sigma > 0$. Specifically, for any configuration $\mathbf{x} = \{(z_1, m_1), \dots, (z_{n(\mathbf{x})}, m_{n(\mathbf{x})})\}$ and any $m \in \mathbb{M}$,

$$k_\beta^m(\mathbf{x}, z) = \frac{1}{n(\mathbf{x})} \sum_{i=1}^{n(\mathbf{x})} \frac{1}{v(z_i, z)} \exp\left(-\frac{\|z - z_i\|^2}{2\sigma^2}\right),$$

where $v(z_i, z) = \int_\Lambda \exp\left(-\frac{\|z - z_i\|^2}{2\sigma^2}\right) dz$. We may easily extend this example by requiring that a particle with label m can only appear close to particles with labels m' .

Example 2 (Birth kernel driven by a potential). In this example the birth

kernel is given through a function $V : E \rightarrow \mathbb{R}$ by :

$$k_\beta^m(\mathbf{x}, z) = \frac{1}{a^m(\mathbf{x})} e^{-(V(\mathbf{x} \cup (z, m)) - V(\mathbf{x}))}$$

where $a^m(\mathbf{x}) = \int_\Lambda e^{-(V(\mathbf{x} \cup (z, m)) - V(\mathbf{x}))} dz$. Given a configuration \mathbf{x} , a new particle is more likely to appear in the vicinity of points $z \in \Lambda$ that make $V(\mathbf{x} \cup (z, m)) - V(\mathbf{x})$ minimal. Following the statistical physics terminology, the function V can then be seen as a potential that the system tends to minimizing at each birth. A typical instance, for $\mathbf{x} = \{(z_1, m_1), \dots, (z_{n(\mathbf{x})}, m_{n(\mathbf{x})})\}$, is $V(\mathbf{x}) = \sum_{i \neq j} \Phi_{m_i, m_j}(z_i - z_j)$ for some pair potential functions $\Phi_{m, m'}$, $m, m' \in \mathbb{M}$, see [LLGM22] for some examples.

Concerning the mutation transition kernel, we set for any $\mathbf{x} = \{x_1, \dots, x_{n(\mathbf{x})}\}$, where $x_i = (z_i, m_i)$, and any $F \in \mathcal{E}$,

$$\nu_\tau(\mathbf{x}, F) = \sum_{i=1}^{n(\mathbf{x})} \sum_{m \in \mathbb{M}} \mathbb{1}_{(\mathbf{x} \setminus (z_i, m_i)) \cup (z_i, m) \in F},$$

so that a transition only concerns a change of label of an existing particle. For the density k_τ , we assume that

$$k_\tau(\mathbf{x}, \mathbf{y}) = \begin{cases} s(x_i, \mathbf{x}) q_m(x_i, \mathbf{x}) & \text{if } \exists x_i \in \mathbf{x}, \exists m \in \mathbb{M}, \mathbf{y} = (\mathbf{x} \setminus (z_i, m_i)) \cup (z_i, m), \\ 0 & \text{otherwise.} \end{cases}$$

In this expression, $s(x_i, \mathbf{x}) \in [0, 1]$ is the probability that the particle x_i in \mathbf{x} changes his label and $q_m(x_i, \mathbf{x})$ is the probability that, given that the particle $x_i \in \mathbf{x}$ mutates, its label changes from m_i to m . We thus have $\sum_{i=1}^{n(\mathbf{x})} s(x_i, \mathbf{x}) = 1$ and $\sum_{m \in \mathbb{M}} q_m(x_i, \mathbf{x}) = 1$ with $q_{m_i}(x_i, \mathbf{x}) = 0$. An example is provided below.

Example 3 (Mutation kernel given by a transition matrix). A natural example is when the particle to modify is uniformly chosen among all particles, and then the change of label is done according to a transition matrix with coefficients $(p_{m, m'})_{m, m' \in \mathbb{M}}$. This corresponds to the choices

$$s(x_i, \mathbf{x}) = \frac{1}{n(\mathbf{x})} \quad \text{and} \quad q_m(x_i, \mathbf{x}) = p_{m_i, m},$$

where $p_{m, m} = 0$ and for all $m \in \mathbb{M}$, $\sum_{m' \in \mathbb{M}} p_{m, m'} = 1$.

Finally, for the death transition kernel, we set for any $\mathbf{x} = \{x_1, \dots, x_{n(\mathbf{x})}\}$ and any $F \in \mathcal{E}$, $\nu_\delta(\mathbf{x}, F) = \sum_{i=1}^{n(\mathbf{x})} \mathbb{1}_{\mathbf{x} \setminus x_i \in F}$, so that this transition only concerns the death of an existing particle. For the density,

$$k_\delta(\mathbf{x}, \mathbf{y}) = \begin{cases} \omega(x_i, x) & \text{if } \mathbf{y} = \mathbf{x} \setminus x_i, \\ 0 & \text{otherwise,} \end{cases}$$

where $\omega(x_i, \mathbf{x}) \in [0, 1]$ is the probability that the particle x_i in \mathbf{x} dies, with $\sum_{i=1}^{n(\mathbf{x})} \omega(x_i, \mathbf{x}) = 1$. A simple example is the uniform death where $\omega(x_i, \mathbf{x}) = 1/n(\mathbf{x})$ for any i .

3.1.3.3 The inter-jumps motion

Remember that between two jumps, the BDMM process $(\mathbf{X}_t)_{t \geq 0}$ has a constant cardinality $n = n(\mathbf{X}_t)$ and constant labels in \mathbb{M} . To specify the inter-jumps motion of $(\mathbf{X}_t)_{t \geq 0}$, it is then enough to define the dynamics of a process $(\mathbf{Y}_t^n)_{t \geq 0}$ on each subspace E_n where

$$\mathbf{Y}_t^n = \{(z_{1,t}, m_1), \dots, (z_{n,t}, m_n)\}.$$

Accordingly, the position and the continuous mark of each particle may move, that is $z_{i,t} \in \Lambda$ may move, but the label m_i will remain constant. To define such dynamics, we first need to come back to a standard system of n ordered (labelled) particles in $(\Lambda \times \mathbb{M})^n$, that is

$$\tilde{\mathbf{Y}}_t^n = ((z_{1,t}, m_1), \dots, (z_{n,t}, m_n)).$$

The full definition of the move process $(\mathbf{Y}_t)_{t \geq 0}$ on E then follows the three steps :

1. Define the dynamics of $\tilde{\mathbf{Y}}_t^n$ on $(\Lambda \times \mathbb{M})^n$ thanks to a system of stochastic differential equations, as (M^n) specified below, where the motion acts only in Λ^n .
2. Provided this system yields a solution whose distribution satisfies the permutation equivariance property (see below), deduce $\mathbf{Y}_t^n = \Pi_n(\tilde{\mathbf{Y}}_t^n)$ on E_n , where Π_n is given by (3.1).
3. Then define $(\mathbf{Y}_t)_{t \geq 0}$ on E by $\mathbf{Y}_t = \sum_{n \geq 0} \mathbf{Y}_t^n \mathbb{1}_{\{\mathbf{Y}_0 \in E_n\}}$.

This construction and the following facts are detailed in [LLGM22]. The equivariance property means that for any permutation π of $\{1, \dots, n\}$, the law of

$\tilde{\mathbf{Y}}_t^n = ((z_{1,t}, m_1), \dots, (z_{n,t}, m_n))$ given $\tilde{\mathbf{Y}}_0^n = ((z_{\pi(1)}, m_{\pi(1)}), \dots, (z_{\pi(n)}, m_{\pi(n)}))$ is the

same as the distribution of $((z_{\pi(1),t}, m_{\pi(1)}), \dots, (z_{\pi(n),t}, m_{\pi(n)}))$ given $\tilde{\mathbf{Y}}_0^{|n} = ((z_1, m_1), \dots, (z_n, m_n))$. This property says that if we rearrange the ordering of the coordinates of the initial state, the ordering of the solution is rearranged in the same way, so that it makes sense to deduce $\mathbf{Y}_t^{|n}$ from $\tilde{\mathbf{Y}}_t^{|n}$. Under this assumption, the Markov process \mathbf{Y} is well-defined on E and its transition kernel $Q_t^{\mathbf{Y}}$, defined for any bounded and measurable function f on E by $Q_t^{\mathbf{Y}} f(x) = \mathbb{E}[f(\mathbf{Y}_t) | \mathbf{Y}_0 = x]$, reads

$$Q_t^{\mathbf{Y}} f(x) = \sum_{n \geq 0} Q_t^{\tilde{\mathbf{Y}}^{|n}} (f \circ \Pi_n)((x_1, \dots, x_n)) \mathbf{1}_{E_n}(x),$$

where $Q_t^{\tilde{\mathbf{Y}}^{|n}}$ denotes the transition kernel of $\tilde{\mathbf{Y}}^{|n}$ in $(\Lambda \times \mathbb{M})^n$. Moreover, \mathbf{Y} is continuous in (E, d_1) whenever for any n , $\tilde{\mathbf{Y}}^{|n}$ is continuous in $(\Lambda \times \mathbb{M})^n$ for the usual Euclidean distance.

As a consequence of this construction, the definition of $(\mathbf{Y}_t)_{t \geq 0}$ boils down to the first step above, that is the specification of the system of SDE $(M^{|n})$, provided the latter is permutation equivariant. We assume in this paper that $\tilde{\mathbf{Y}}_t^{|n} = (\mathbf{Z}_t, \mathbf{m})$, starting at $t = 0$ from (\mathbf{z}, \mathbf{m}) where $\mathbf{z} = (z_1, \dots, z_n)$, is the solution of

$$M^{|n}(\mathbf{z}, \mathbf{m}) : \begin{cases} dz_{i,t} = v_{i,n}(\mathbf{Z}_t, \mathbf{m}) dt + \sigma_{i,n}(\mathbf{Z}_t, \mathbf{m}) dB_{i,t}, & t \geq 0, \quad i = 1, \dots, n, \\ z_{i,0} = z_i, & i = 1, \dots, n. \end{cases}$$

Here the drift functions $v_{1,n}, \dots, v_{n,n}$ take their values in \mathbb{R}^d , the diffusion coefficients $\sigma_{1,n}, \dots, \sigma_{n,n}$ are (d, d) invertible matrices and $(B_{1,t})_{t \geq 0}, \dots, (B_{n,t})_{t \geq 0}$ are independent standard Brownian motions on \mathbb{R}^d . We assume that the functions $v_{i,n}$ and $\sigma_{i,n}$ are globally Lipschitzian, so that a strong solution exists. If $\Lambda \neq \mathbb{R}^d$, we further assume that edge conditions (reflective or periodic) are added to ensure that the solution stays in $(\Lambda \times \mathbb{M})^n$, see for instance [FG07].

The permutation equivariance property is ensured if for any $i = 1, \dots, n$ and any $\pi \in S_n$,

$$\begin{cases} v_{i,n}((z_{\pi(1)}, m_{\pi(1)}), \dots, (z_{\pi(n)}, m_{\pi(n)})) = v_{\pi(i),n}((z_1, m_1), \dots, (z_n, m_n)), \\ \sigma_{i,n}((z_{\pi(1)}, m_{\pi(1)}), \dots, (z_{\pi(n)}, m_{\pi(n)})) = \sigma_{\pi(i),n}((z_1, m_1), \dots, (z_n, m_n)). \end{cases}$$

This is for instance the case if $v_{i,n}(\mathbf{Z}, \mathbf{m}) = v(z_i, m_i)$ and $\sigma_{i,n}(\mathbf{Z}, \mathbf{m}) = \sigma(z_i, m_i)$ for some functions v and σ , a situation where there is no interaction between the particles, as considered in the microbiological applications in [BVK19] and [BLSK23].

An example that includes interactions through a potential function is provided in the following example.

Example 4 (Langevin diffusion). In this example the motion of each particle depends on an interaction force with the other particles, driven by a pair potential functions $\Phi_{m,m'}$, $m, m' \in \mathbb{M}$, as in Example 2. Specifically, in this model $\sigma_i(\mathbf{Z}, \mathbf{m}) = \sigma_{m_i}$ and

$$v_{i,n}(\mathbf{Z}, \mathbf{m}) = - \sum_{j \neq i} \Phi_{m_i, m_j}(z_i - z_j).$$

For existence, each pair potential $\Phi_{m,m'}$ must be smooth enough. We refer to [FG07] and [LLGM22] for details and examples.

For later purposes, let us specify the likelihood of $\mathbf{Y}_t^{|n}$ constructed as above. Let us first rewrite $(M^{|n})$ so that it takes the form of a standard stochastic differential equation. Denote by $\bar{\sigma}_n$ the block diagonal matrix of size (nd, nd) formed by the $\sigma_{i,n}$ matrices, and by \bar{v}_n the vector of size nd formed by the concatenation of the $v_{i,n}$ vectors. Then, by denoting B_t the vector formed by the concatenation of the vectors $B_{i,t}$, we have

$$M^{|n}(\mathbf{z}, \mathbf{m}) : \begin{cases} d\mathbf{Z}_t = \bar{v}_n(\mathbf{Z}_t, \mathbf{m})dt + \bar{\sigma}_n(\mathbf{Z}_t, \mathbf{m})dB_t, & t \geq 0, \\ \mathbf{Z}_0 = \mathbf{z}. \end{cases} \quad (3.4)$$

It is well known, cf [LS13, Kut04], that the Radon-Nikodym density of the solution $(\mathbf{Z}_t)_{t \geq 0}$ of $M^{|n}(\mathbf{z}, \mathbf{m})$ with respect to the reference process $\mathbf{U}_t = \mathbf{z} + \int_0^t \bar{\sigma}_n(\mathbf{U}_s, \mathbf{m})dB_s$, on the interval $[0, t]$ reads

$$L(\mathbf{Z}_{[0,t]}, \mathbf{m}) = \exp \left(\int_0^t \bar{v}_n(\mathbf{Z}_s, \mathbf{m})^T a_n^{-1}(\mathbf{Z}_s, \mathbf{m}) d\mathbf{Z}_s - \frac{1}{2} \int_0^t \bar{v}_n(\mathbf{Z}_s, \mathbf{m})^T a_n^{-1}(\mathbf{Z}_s, \mathbf{m}) \bar{v}_n(\mathbf{Z}_s, \mathbf{m}) ds \right), \quad (3.5)$$

where $a_n(\mathbf{z}, \mathbf{m}) = \bar{\sigma}_n(\mathbf{z}, \mathbf{m})\bar{\sigma}_n^T(\mathbf{z}, \mathbf{m})$. This gives us the density on $[0, t[$ of the process $\tilde{\mathbf{Y}}_t^{|n} = (\mathbf{Z}_t, \mathbf{m})$ with respect to the reference process $(\mathbf{U}_t, \mathbf{m})_{t \geq 0}$ that belongs to $(\Lambda \times \mathbb{M})^n$. Finally, by our construction above, if $(\mathbf{Z}_t, \mathbf{m})$ is such that $\mathbf{Y}_t^{|n} = \Pi_n((\mathbf{Z}_t, \mathbf{m}))$, then the density of $\mathbf{Y}_t^{|n}$ with respect to $\Pi_n((\mathbf{U}_t, \mathbf{m}))$ on $[0, t]$ takes the same form and does not depend on the ordering chosen to define \mathbf{Z}_t, \mathbf{z} and \mathbf{m} from $\mathbf{Y}_t^{|n}$, thanks to the permutation equivariance property.

3.2 LAN property

3.2.1 Likelihood of the BDMM process

As described in the previous section, the BDMM process $(\mathbf{X}_t)_{t \geq 0}$ depends on the intensity functions β , δ , τ , assumed to be continuous on E , on kernel densities for the births, the deaths and the mutations, denoted by $k_\beta(\mathbf{x}, \mathbf{y})$, $k_\delta(\mathbf{x}, \mathbf{y})$ and $k_\tau(\mathbf{x}, \mathbf{y})$, as detailed in Section 3.1.3.2, and on a continuous Markov diffusion model on E that drives the inter-jump motion, see Section 3.1.3.3.

Remember that between two jumps T_i and T_{i+1} , $\mathbf{X}_t = \mathbf{Y}_{t-T_i}^{(i)}$ where $\mathbf{Y}^{(i)}$ has the same law as \mathbf{Y} . Write $\mathbf{Y}_t^{(i)} = \{(\mathbf{Z}_t^{(i)}, \mathbf{m}^{(i)})\}$ where $\mathbf{m}^{(i)}$ does not depend on t , as required for the labels of the continuous Markov process that drives the inter-jumps motion. As assumed in Section 3.1.3.3, $\mathbf{Z}^{(i)}$ is the solution of $M^{|n}(\mathbf{z}^{(i)}, \mathbf{m}^{(i)})$ given by (3.4), where $(\mathbf{z}^{(i)}, \mathbf{m}^{(i)})$ are such that $\mathbf{X}_{T_i} = \{(\mathbf{z}^{(i)}, \mathbf{m}^{(i)})\}$ and $n = n(\mathbf{X}_{T_i})$. The likelihood of the inter-jump motion between T_i and T_{i+1} , given T_i , T_{i+1} and \mathbf{X}_{T_i} , is thus $L(\mathbf{Z}_{[0, T_{i+1}-T_i]}^{(i)}, \mathbf{m}^{(i)})$ given by (3.5). For simplicity we will simply write the latter $L(\mathbf{X}_{[T_i, T_{i+1}[})$ in the following, since $\mathbf{X}_t = \{(\mathbf{Z}_{t-T_i}^{(i)}, \mathbf{m}^{(i)})\}$ for all $t \in [T_i, T_{i+1}[$.

Assume that all features of the BDMM process depend on a parameter $\vartheta \in \Theta \subset \mathbb{R}^\ell$ for a given $\ell > 0$. To emphasize this dependence, we introduce the argument ϑ into the notation of each of them. For example we will write $\beta(\vartheta, \mathbf{x})$ instead of $\beta(\mathbf{x})$, $k_\beta(\vartheta, \mathbf{x}, \mathbf{y})$ instead of $k_\beta(\mathbf{x}, \mathbf{y})$, $L(\vartheta, \mathbf{X}_{[T_i, T_{i+1}[})$ instead of $L(\mathbf{X}_{[T_i, T_{i+1}[})$, and so on. Note however that in (3.5), ϑ only appears in the drift coefficient \bar{v}_n , i.e. $\bar{v}_n(\mathbf{X}_s) = \bar{v}_n(\vartheta, \mathbf{X}_s)$, but not in the diffusion coefficient $\bar{\sigma}_n$ which is part of the reference process. This is indeed a common fact that for continuous time observations, the diffusion coefficient can be considered to be known [Kut04], see also Appendix B.

Let us denote the set of birth times up to time t by

$$\mathcal{B}_t = \{i \in \mathbb{N}^*, T_i \leq t \text{ and } T_i \text{ is a birth time}\},$$

and similarly \mathcal{D}_t and \mathcal{T}_t for the set of death times and mutation times up to time t .

The following theorem provides the likelihood of the BDMM process. When there is no move and no mutation, we recover the formula given in [MS94] for their applications to the dynamics of dunes. It also has a consistent expression with the likelihood of a system of particles with killing and jumps, as derived in [Löc02b]. The justifications, detailed in Appendix C, are however different and simpler, as

they only rely on successive backward applications of conditional expectations.

Theorem 1. *Let $(\mathbf{X}_t)_{t \geq 0}$ be a BDMM parameterized by $\vartheta \in \Theta$, according to the above conditions, and observed in continuous time on $[0, t]$, for $t > 0$. Then, the likelihood $\mathcal{L}_t(\vartheta)$ is expressed as follows :*

$$\mathcal{L}_t(\vartheta) = \mathcal{L}_t^{\mathcal{B}}(\vartheta) \mathcal{L}_t^{\mathcal{D}}(\vartheta) \mathcal{L}_t^{\mathcal{T}}(\vartheta) \mathcal{L}_t^{\text{move}}(\vartheta),$$

with

$$\begin{aligned} \mathcal{L}_t^{\mathcal{B}}(\vartheta) &= \exp\left(-\int_0^t \beta(\vartheta, \mathbf{X}_s) ds\right) \prod_{i \in \mathcal{B}_t} k_\beta(\vartheta, \mathbf{X}_{T_i^-}, \mathbf{X}_{T_i}) \beta(\vartheta, \mathbf{X}_{T_i^-}), \\ \mathcal{L}_t^{\mathcal{D}}(\vartheta) &= \exp\left(-\int_0^t \delta(\vartheta, \mathbf{X}_s) ds\right) \prod_{i \in \mathcal{D}_t} k_\delta(\vartheta, \mathbf{X}_{T_i^-}, \mathbf{X}_{T_i}) \delta(\vartheta, \mathbf{X}_{T_i^-}), \\ \mathcal{L}_t^{\mathcal{T}}(\vartheta) &= \exp\left(-\int_0^t \tau(\vartheta, \mathbf{X}_s) ds\right) \prod_{i \in \mathcal{T}_t} k_\tau(\vartheta, \mathbf{X}_{T_i^-}, \mathbf{X}_{T_i}) \tau(\vartheta, \mathbf{X}_{T_i^-}), \\ \mathcal{L}_t^{\text{move}}(\vartheta) &= L(\vartheta, \mathbf{X}_{[T_{N_t}, t]}) \prod_{i=0}^{N_t-1} L(\vartheta, \mathbf{X}_{[T_i, T_{i+1}]}). \end{aligned}$$

Note that due to the Markov nature of the process, the likelihood is a simple product. As a consequence, inference on a specific aspect of the process can be done independently on the other characteristics, provided the parameter is different in each characteristic. As an illustration, we will infer the birth kernel k_β in our application in Chapter 4, without needing to specify or estimate the other terms.

3.2.2 LAN

Recall that all the features of the BDMM process depend on a parameter $\vartheta \in \Theta \subset \mathbb{R}^\ell$. In the following, for a multivariate function f , we denote by $\nabla_i f$ the gradient (or Jacobian matrix if f is multivariate) with respect to the i -th variable. We will also denote by $B(\vartheta, \rho)$ the ball centred at ϑ with radius $\rho > 0$.

We list below the assumptions we make to obtain the local asymptotic normality of our model. They are of two kinds : the first one **(H)** ensures that the process is non-explosive and geometric ergodic. Conditions implying this hypothesis concern the intensity functions : it basically holds if the total intensity function α is bounded and the death intensity function δ sufficiently compensates for the birth intensity β . Precise conditions are given in Appendix E, following [LLGM22]. The other as-

assumptions are regularity hypotheses on the way each feature of the BDMM process is parametrised. Although quite technical at first sight, we show in Section 3.2.3 that they are satisfied in many situations, including the examples of Section 3.1.3.

(H) The process is non-explosive, i.e. $N_t < \infty$ for any $t \geq 0$, and there exists a measure μ_∞^ϑ on E , $r_\vartheta > 0$ and $c_\vartheta : E \rightarrow \mathbb{R}_+^*$ such that for any $t > 0$, any $\mathbf{x}_0 \in E$ and any measurable and bounded function g on E ,

$$\left| \frac{1}{t} \mathbb{E}_{\mathbf{x}_0}^\vartheta \left(\int_0^t g(\mathbf{X}_s) ds \right) - \int_E g(\mathbf{y}) \mu_\infty^\vartheta(d\mathbf{y}) \right| \leq \frac{c_\vartheta(\mathbf{x}_0) \|g\|_\infty}{r_\vartheta t}.$$

Concerning the intensity function γ (whether γ is β , δ or τ), we assume that it is differentiable with respect to ϑ and the two following hypotheses.

A(γ) For all $\vartheta \in \Theta$, $\mathbf{x} \mapsto \frac{\|\nabla_1 \gamma(\vartheta, \mathbf{x})\|^2}{\gamma(\vartheta, \mathbf{x})} \mathbb{1}_{\gamma(\vartheta, \mathbf{x}) > 0}$ is bounded.

B(γ) For all $\vartheta \in \Theta$, $\exists \rho(\vartheta) > 0$ such that $\mathbf{x} \mapsto f_\vartheta^\gamma(\mathbf{x}, \rho(\vartheta)) \gamma(\vartheta, \mathbf{x})$ is bounded, where

$$f_\vartheta^\gamma(\mathbf{x}, \rho) = \sup_{u \in B(\vartheta, \rho)} \left(\|\nabla_1 \gamma(u, \mathbf{x}) - \nabla_1 \gamma(\vartheta, \mathbf{x})\|^2 \right) \frac{\mathbb{1}_{\gamma(\vartheta, \mathbf{x}) > 0}}{(\gamma(\vartheta, \mathbf{x}))^2}. \quad (3.6)$$

Similarly, we assume that the transition kernel k_γ is differentiable with respect to ϑ and the two following hypotheses.

A(k_γ, γ) For all $\vartheta \in \Theta$, $\mathbf{x} \mapsto \gamma(\vartheta, \mathbf{x}) \int_E \frac{\|\nabla_1 k_\gamma(\vartheta, \mathbf{x}, \mathbf{y})\|^2}{k_\gamma(\vartheta, \mathbf{x}, \mathbf{y})} \mathbb{1}_{k_\gamma(\vartheta, \mathbf{x}, \mathbf{y}) > 0} \nu_\gamma(\mathbf{x}, d\mathbf{y})$ is bounded.

B(k_γ, γ) For all $\vartheta \in \Theta$, $\exists \rho(\vartheta) > 0$ such that $\mathbf{x} \mapsto f_\vartheta^{k_\gamma}(\mathbf{x}, \rho(\vartheta)) \gamma(\vartheta, \mathbf{x})$ is bounded, where

$$f_\vartheta^{k_\gamma}(\mathbf{x}, \rho) = \int_E \sup_{u \in B(\vartheta, \rho)} \left(\|\nabla_1 k_\gamma(u, \mathbf{x}, \mathbf{y}) - \nabla_1 k_\gamma(\vartheta, \mathbf{x}, \mathbf{y})\|^2 \right) \frac{\mathbb{1}_{k_\gamma(\vartheta, \mathbf{x}, \mathbf{y}) > 0}}{k_\gamma(\vartheta, \mathbf{x}, \mathbf{y})} \nu_\gamma(\mathbf{x}, d\mathbf{y}). \quad (3.7)$$

Finally, concerning the inter-jump motion given as a solution of (3.4), we assume that its drift function \bar{v}_n is differentiable with respect to ϑ and :

A_{move} For all $n \geq 1$, $\forall \vartheta \in \Theta$, $\mathbf{x} \mapsto \|\bar{\sigma}_n(\mathbf{x})^{-1} \nabla_1 \bar{v}_n(\vartheta, \mathbf{x})\|$ is bounded.

\mathbf{B}_{move} For all $n \geq 1$, $\forall \vartheta \in \Theta$, $\exists \rho(\vartheta) > 0$ such that $\mathbf{x} \mapsto \|\bar{\sigma}_n(\mathbf{x})^{-1}\|^2 f_{\vartheta}^{\text{move}}(\mathbf{x}, \rho(\vartheta))$ is bounded, where

$$f_{\vartheta}^{\text{move}}(\mathbf{x}, \rho) = \sup_{u \in B(\vartheta, \rho)} \|\nabla_1 \bar{v}_n(u, \mathbf{x}) - \nabla_1 \bar{v}_n(\vartheta, \mathbf{x})\|^2. \quad (3.8)$$

In order to state the LAN property, we introduce, for any $\vartheta \in \Theta$ and $N \in \mathbb{N}^*$, the following \mathcal{F}_{tN} -martingales, along with their associated angle brackets. The fact that they are truly martingales and the expression of their brackets are justified in the proof of the following theorem deferred to Appendix D. We recall that $a_n = \bar{\sigma}_n \bar{\sigma}_n^T$.

$$M_{\gamma}^{N, \vartheta}(t) = \frac{1}{\sqrt{N}} \int_0^{tN} \frac{\nabla_1 \gamma(\vartheta, \mathbf{X}_{s-})}{\gamma(\vartheta, \mathbf{X}_{s-})} dM_s^{\gamma}, \quad (3.9)$$

$$M_{k_{\gamma}}^{N, \vartheta}(t) = \frac{1}{\sqrt{N}} \left(\int_0^{tN} \frac{\nabla_1 k_{\gamma}(\vartheta, \mathbf{X}_{s-}, \mathbf{X}_s)}{k_{\gamma}(\vartheta, \mathbf{X}_{s-}, \mathbf{X}_s)} dN_s^{\gamma} - \int_0^{tN} \int_E \nabla_1 k_{\gamma}(\vartheta, \mathbf{X}_s, \mathbf{y}) \nu_{\gamma}(\mathbf{X}_s, d\mathbf{y}) \gamma(\vartheta, \mathbf{X}_s) ds \right), \quad (3.10)$$

$$M_L^{N, \vartheta}(t) = \frac{1}{\sqrt{N}} \sum_{i \geq 0} \int_0^{tN} \mathbf{1}_{[T_i, T_{i+1}[}(s) (\nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_{s-}))^T a_n^{-1}(\mathbf{X}_{s-}) dM_{s-T_i}^{(i)}, \quad (3.11)$$

where $M_t^{\gamma} = N_t^{\gamma} - \int_0^t \gamma(\vartheta, \mathbf{X}_s) ds$ and $M_s^{(i)} = (\mathbf{Z}_s^{(i)} - \mathbf{Z}_0^{(i)} - \int_0^s \bar{v}_n(\vartheta, \mathbf{Z}_u^{(i)}) du) \mathbf{1}_{s \geq T_i}$ if we agree that $\mathbf{X}_t = \{(\mathbf{Z}_{t-T_i}^{(i)}, \mathbf{m}^{(i)})\}$ when $t \in [T_i, T_{i+1}[$. Their associated angle brackets are given by

$$\langle M_{\gamma}^{N, \vartheta} \rangle(t) = \frac{1}{N} \int_0^{tN} \mathbf{1}_{\gamma > 0} \frac{(\nabla_1 \gamma)(\nabla_1 \gamma)^T}{\gamma}(\vartheta, \mathbf{X}_s) ds, \quad (3.12)$$

$$\langle M_{k_{\gamma}}^{N, \vartheta} \rangle(t) = \frac{1}{N} \int_0^{tN} \int_E \mathbf{1}_{k_{\gamma} > 0} \frac{\nabla_1 k_{\gamma}(\nabla_1 k_{\gamma})^T}{k_{\gamma}}(\vartheta, \mathbf{X}_s, \mathbf{y}) \nu_{\gamma}(\mathbf{X}_s, d\mathbf{y}) \gamma(\vartheta, \mathbf{X}_s) ds, \quad (3.13)$$

$$\langle M_L^{N, \vartheta} \rangle(t) = \frac{1}{N} \int_0^{tN} (\nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s))^T a_n^{-1}(\mathbf{X}_s) \nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s) ds. \quad (3.14)$$

Finally denoting $\mathcal{S} = \{L, \beta, \delta, \tau, k_{\beta}, k_{\delta}, k_{\tau}\}$, we introduce

$$M^{N, \vartheta}(t) = \sum_{i \in \mathcal{S}} M_i^{N, \vartheta}(t), \quad (3.15)$$

the angle brackets of which reads

$$\langle M^{N, \vartheta} \rangle(t) = \sum_{i \in \mathcal{S}} \langle M_i^{N, \vartheta} \rangle(t). \quad (3.16)$$

Theorem 2. *Under the same conditions as in Theorem 1, suppose that (\mathbf{H}) , $\mathbf{A}(\gamma)$, $\mathbf{B}(\gamma)$, $\mathbf{A}(k_\gamma, \gamma)$, $\mathbf{B}(k_\gamma, \gamma)$, \mathbf{A}_{move} and \mathbf{B}_{move} are satisfied (whether γ is β , δ or τ). Then for all $\vartheta \in \mathbb{R}^\ell$, $t > 0$, $h \in \mathbb{R}^\ell$ and $N \in \mathbb{N}^*$, we have the following decomposition :*

$$\log \frac{\mathcal{L}_{tN}(\vartheta + \frac{h}{\sqrt{N}})}{\mathcal{L}_{tN}(\vartheta)} = h^T M^{N,\vartheta}(t) - \frac{1}{2} h^T \langle M^{N,\vartheta} \rangle(t) h + \text{Rem}_{N,\vartheta,h}(t), \quad (3.17)$$

where $\text{Rem}_{N,\vartheta,h}(t) \rightarrow 0$ in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability when $N \rightarrow +\infty$, for any $\mathbf{x}_0 \in E$, and where $M^{N,\vartheta}$ and $\langle M^{N,\vartheta} \rangle$ are given by (3.15) and (3.16). Moreover, for all $t > 0$ and $\vartheta \in \Theta$,

$$(M^{N,\vartheta}(t), \langle M^{N,\vartheta} \rangle(t)) \xrightarrow{N \rightarrow +\infty} (\mathcal{N}(0, tJ(\vartheta)), tJ(\vartheta)), \quad (3.18)$$

where the joint convergence takes place in distribution and where the matrix $J(\vartheta)$, of size (ℓ, ℓ) , has for expression :

$$\begin{aligned} J(\vartheta) &= \lim_{N \rightarrow +\infty} \frac{1}{t} \langle M^{N,\vartheta} \rangle(t) \\ &= \int_E (\nabla_1 \bar{v}_n(\vartheta, \mathbf{x}))^T a_n^{-1}(\mathbf{x}) \nabla_1 \bar{v}_n(\vartheta, \mathbf{x}) \mu_\infty^\vartheta(d\mathbf{x}) \\ &\quad + \sum_{\gamma \in \{\beta, \delta, \tau\}} \int_E \mathbf{1}_{\gamma > 0} \frac{(\nabla_1 \gamma)(\nabla_1 \gamma)^T}{\gamma}(\vartheta, \mathbf{x}) \mu_\infty^\vartheta(d\mathbf{x}) \\ &\quad + \sum_{\gamma \in \{\beta, \delta, \tau\}} \int_E \int_E \mathbf{1}_{k_\gamma > 0} \frac{\nabla_1 k_\gamma (\nabla_1 k_\gamma)^T}{k_\gamma}(\vartheta, \mathbf{x}, \mathbf{y}) \nu_\gamma(\mathbf{x}, d\mathbf{y}) \gamma(\vartheta, \mathbf{x}) \mu_\infty^\vartheta(d\mathbf{x}). \end{aligned} \quad (3.19)$$

Under additional regularity assumptions on the model, as detailed for instance in [Ibr13], the LAN property implies that the maximum likelihood estimator is asymptotically efficient and asymptotically Gaussian with covariance $J(\vartheta)^{-1}$. The following corollary summarise this consequence, the proof of which may be found in [Ibr13].

Corollary 1. *In addition to the setting of Theorem 2, assume that the regularity conditions N2-N4 of [Ibr13, Chapter III] are satisfied. Then the maximum likelihood estimator $\hat{\vartheta}_N$, defined by $\mathcal{L}_N(\hat{\vartheta}_N) = \sup_\vartheta \mathcal{L}_N(\vartheta)$, where \mathcal{L}_N is given in Theorem 1, is asymptotically efficient and satisfies*

$$\sqrt{N}(\hat{\vartheta}_N - \vartheta) \xrightarrow{N \rightarrow +\infty} \mathcal{N}(0, J(\vartheta)^{-1}),$$

where the convergence is in distribution and $J(\vartheta)$ is given by (3.19).

Note that in Theorem 2, the length of the observed trajectory is Nt , and the asymptotic regime stands for N large and any $t > 0$. We set in the above corollary $t = 1$ for simplicity and considered a trajectory of size N .

In practice, the asymptotic covariance matrix $J(\vartheta)^{-1}$ is estimated by $J_N(\hat{\vartheta}_N)^{-1}$, where $\hat{\vartheta}_N$ is the maximum likelihood estimator and J_N is the empirical version of J given by (3.19), viz.

$$J_N(\vartheta) = \langle M^{N,\vartheta} \rangle(1),$$

where $\langle M^{N,\vartheta} \rangle$ is given by (3.16) and involves the terms (3.12), (3.13) and (3.14). All these quantities are computable from a single trajectory over $[0, N]$, as implemented in Section 2.1.2.

3.2.3 Examples

We show in this section that the regularity conditions $\mathbf{A}(\gamma)$, $\mathbf{B}(\gamma)$, $\mathbf{A}(k_\gamma, \gamma)$, $\mathbf{B}(k_\gamma, \gamma)$, \mathbf{A}_{move} and \mathbf{B}_{move} are satisfied under mild assumptions for the examples considered in Section 3.1.3.

For the intensity functions, whether $\gamma(\mathbf{x}) = \gamma$ is constant or $\gamma(\mathbf{x}) = \gamma n(\mathbf{x})$, for $\gamma > 0$, and the parameter of interest is $\vartheta = \gamma$, we have $f_\vartheta^\gamma = 0$ in (3.6), so that $\mathbf{B}(\gamma)$ is obviously satisfied. For $\mathbf{A}(\gamma)$, it clearly holds in the first case, while it is true in the second case if we assume that there is a maximal number of particles, i.e. $\sup_{\mathbf{x}} n(\mathbf{x}) < n^*$ for some $n^* > 0$. Note that the latter restriction is purely theoretical, but it is not a limitation in practice. It is also an assumption made in [LLG21] and it amounts to the specific setting of condition (E.2) presented in appendix, that implies geometric ergodicity of the process, i.e. **(H)**.

Concerning the transition kernels, the parameters of interest are typically involved in the birth kernel and/or the mutation kernel. For the birth kernel, remember that it reads as in (3.3), i.e. $k_\beta(\mathbf{x}, \mathbf{y}) = p_m k_\beta^m(\mathbf{x}, z)$ and let us examine Examples 1 and 2 for k_β^m .

Example 1 (continued). In this example k_β^m is a mixture of Gaussian distributions with deviation $\sigma > 0$, so that the parameters of interest for the birth kernel are in this case p_m , $m \in \mathbb{M}$, and σ . It is not difficult to check that if Λ is a bounded set, then all partial derivatives of $k_\beta(\mathbf{x}, \mathbf{y})$ with respect to each parameter are uniformly

bounded in \mathbf{x} and \mathbf{y} . On the other hand $k_\beta(\mathbf{x}, \mathbf{y})$ itself is lower bounded in \mathbf{x} and \mathbf{y} under the same setting. Consequently, $\mathbf{A}(k_\gamma, \gamma)$ holds true whenever Λ is a bounded set and the birth intensity $\beta(\mathbf{x})$ is bounded in \mathbf{x} . The latter is in particular true if β satisfies the setting discussed in the previous paragraph. For $\mathbf{B}(k_\gamma, \gamma)$, note that a rough control consists in upper-bounding the squared norm of the difference in (3.7) by twice the sum of each squared norm. Then the same bounds as previously can be used to check $\mathbf{B}(k_\gamma, \gamma)$ in the same setting.

Example 2 (continued). Assume that

$$k_\beta^m(\mathbf{x}, z) = \frac{e^{-\vartheta'_m S_m(\mathbf{x}, z)}}{a^m(\mathbf{x})},$$

where $\vartheta_m = (\vartheta_{m, m'})_{m' \in \mathbb{M}}$, $a^m(\mathbf{x}) = \int_\Lambda e^{-\vartheta'_m S_m(\mathbf{x}, z)} dz$ and S_m is the vector

$$S_m(\mathbf{x}, z) = \left(\sum_{\substack{x_i=(z_i, m_i) \in \mathbf{x} \\ m_i=m'}} \Phi_{m, m'}(z - z_i) \right)_{m' \in \mathbb{M}}.$$

This corresponds to a potential of the form $V(\mathbf{x}) = \sum_{i \neq j} \vartheta_{m_i, m_j} \Phi_{m_i, m_j}(z_i - z_j)$ in Example 2. The parameters of interest here are all the $\vartheta_{m, m'}$, $m, m' \in \mathbb{M}$. Assume that for all $m, m' \in \mathbb{M}$, $\Phi_{m, m'}$ is bounded by L . Then if we denote by $n_{m'}(\mathbf{x})$ the number of particles in \mathbf{x} having label m' , we have

$$\|S_m(\mathbf{x}, z)\|^2 \leq \sum_{m' \in \mathbb{M}} L^2 n_{m'}^2(\mathbf{x}) \leq L^2 n(\mathbf{x}) \sum_{m' \in \mathbb{M}} n_{m'}(\mathbf{x}) = L^2 n^2(\mathbf{x}).$$

Assume in addition that there is a maximal number of particles n^* and that Λ is a bounded set, then using the previous upper-bound, we may prove by elementary inequalities that both $\mathbf{A}(k_\gamma, \gamma)$ and $\mathbf{B}(k_\gamma, \gamma)$ are satisfied.

Concerning the mutation kernel, let us inspect Example 3.

Example 3 (continued). Remember that for this example we have

$$k_\tau(\vartheta, \mathbf{x}, \mathbf{y}) = \frac{1}{n(\mathbf{x})} p_{m_i, m}$$

if there exists $x_i \in \mathbf{x}$ and $m \in \mathbb{M}$ such that $\mathbf{y} = (\mathbf{x} \setminus (z_i, m_i)) \cup (z_i, m)$. The parameter of interest ϑ in this example correspond to the non-null entries $p_{m, m'}$ of the

transition matrix. Denote by p_* the minimal value of these non-null entries. We have $\|\nabla_1 k_\tau(\vartheta, \mathbf{x}, \mathbf{y})\|^2 = 1/n(\mathbf{x})^2$, so that

$$\begin{aligned} & \int_E \frac{\|\nabla_1 k_\tau(\vartheta, \mathbf{x}, \mathbf{y})\|^2}{k_\tau(\vartheta, \mathbf{x}, \mathbf{y})} \mathbb{1}_{k_\tau(\vartheta, \mathbf{x}, \mathbf{y}) > 0} \nu_\tau(\mathbf{x}, d\mathbf{y}) \\ &= \int_E \frac{\|\nabla_1 k_\tau(\vartheta, \mathbf{x}, \mathbf{y})\|^2}{k_\tau(\vartheta, \mathbf{x}, \mathbf{y})^2} k_\tau(\vartheta, \mathbf{x}, \mathbf{y}) \mathbb{1}_{k_\tau(\vartheta, \mathbf{x}, \mathbf{y}) > 0} \nu_\tau(\mathbf{x}, d\mathbf{y}) \\ &\leq \frac{1}{p_*^2} \int_E k_\tau(\vartheta, \mathbf{x}, \mathbf{y}) \nu_\tau(\mathbf{x}, d\mathbf{y}) = \frac{1}{p_*^2}. \end{aligned}$$

Consequently $\mathbf{A}(k_\gamma, \gamma)$ (where $\gamma = \tau$) is satisfied in this setting whenever $\tau(\mathbf{x})$ is bounded in \mathbf{x} , which holds true in the setting discussed in the first paragraph of this section. On the other hand, since $\nabla_1 k_\tau(\vartheta, \mathbf{x}, \mathbf{y})$ does not depend on ϑ , $f_\vartheta^{k_\tau}(\mathbf{x}, \rho) = 0$ in (3.7), so that $\mathbf{B}(k_\gamma, \gamma)$ trivially holds.

Concerning the inter-jump motion and assumptions \mathbf{A}_{move} and \mathbf{B}_{move} , we consider as an example the Langevin diffusion introduced in Example 4.

Example 4 (continued). Assume the same kind of parametrisation of the potential as in Example 2 above, that is $v_{i,n}(\mathbf{Z}, \mathbf{m}) = -\sum_{j \neq i} \vartheta_{m_i, m_j} \Phi_{m_i, m_j}(z_i - z_j)$ in (3.4), where the parameters of interest are again all the $\vartheta_{m, m'}$, $m, m' \in \mathbb{M}$. Remember that $\sigma_i(\mathbf{Z}, \mathbf{m}) = \sigma_{m_i}$ is known. Since the drift function is linear in the parameters, it is easily seen that \mathbf{A}_{move} is satisfied if $\Phi_{m, m'}$ is bounded for any $m, m' \in \mathbb{M}$ and there is a maximal number of particles n^* . As to \mathbf{B}_{move} , we clearly have $f_\vartheta^{\text{move}}(\mathbf{x}, \rho) = 0$ in (3.8), so that this condition holds trivially true.

3.3 Conclusion

In this chapter, we have shown, after giving its likelihood expression, that the BDMM process possesses, under certain regularity assumptions on the various ingredients that characterize the process, the property of local asymptotic normality (LAN). This property is useful to prove the asymptotic normality of the maximum likelihood estimator and to derive an explicit formula for its covariance matrix. Finally, we have shown that these assumptions are satisfied on a large number of examples, in particular those we will need in the next chapter, which concerns the study of real data.

ANALYSIS OF THE COLOCATION PHENOMENON ON SIMULATED AND REAL DATA.

This chapter constitutes the second part of the submitted article [BL24]. Supplementary materials, including the Python code for simulation, the raw data and typical simulated sequences, are available in our online GitHub repository at <https://github.com/balsollier-lisa/Parametric-estimation-of-the-BDMM-process>.

First, a simulation study is carried out, demonstrating the performances of the maximum likelihood estimator of several model parameters, including the estimation of confidence regions. The setting is close to that of the dataset analysed subsequently, in order to provide certain guarantees related the reliability of conclusions. This dataset is a video sequence depicting Langerin proteins and Rab-11 proteins in a living cell, acquired by fluorescence microscopy [BGM⁺14]. As detailed in Section 4.2, particle dynamics are consistent with the realization of BDMM process. Leveraging on this model and its estimation, we quantify the colocalization level between Langerin and Rab-11 proteins.

4.1 Simulation study

In this section, we simulate several trajectories of the BDMM process, for characteristics explained below, and we evaluate the quality of estimation of two target parameters involved in the birth kernel. The chosen characteristics are motivated by the application that we will conduct in the next section, and are the same as those chosen in Section 2.1.2.1 of Chapter 2. We recall them below without precisely detailing each value.

In our context, the particles do not possess continuous marks, but can be of 6

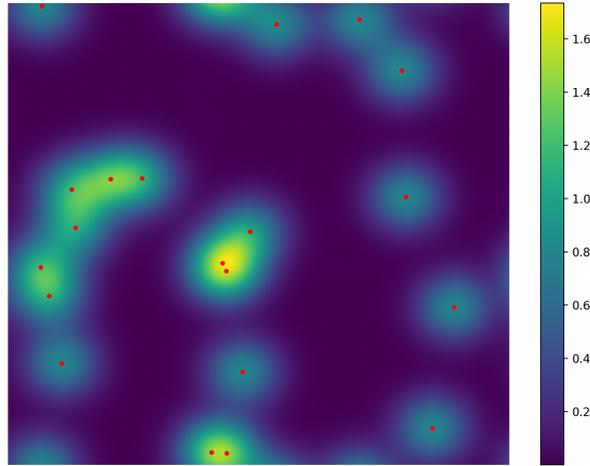


FIGURE 4.1 – Example of the density (4.1) for the birth of a new Langerin particle, given Rab-11 particles already present in the same area (shown as red dots).

different types, which correspond on one hand to two distinct particle types (playing the role of Rab-11 type proteins and Langerin type proteins in the subsequent application), and on the other hand to three possible motion regimes : Brownian, sub-diffusive (according to an Ornstein-Uhlenbeck process), or super-diffusive (according to a Brownian motion with linear drift). Thus, the space \mathbb{M} contains six different labels, but mutations are exclusively possible between motion regimes and not between particles type (see below).

We assume that the movement of each particle is independent of the others and corresponds to one of the three aforementioned regimes. Birth rates are constant and differ only according to the Langerin or Rab-11 type. Death rates (also distinct for Langerin and Rab-11) are proportional to the number of particles present in the cell. Mutations correspond to a transition between two different motion type ; mutation intensity is also assumed to be constant. Regarding transitions, we assume that the death kernel is uniform : each particle has the same probability of disappearing when a death event occurs. For mutations, we follow the setting of Example 3 of Chapter 3, i.e., these occur for any particle uniformly, according to a fixed transition matrix. Finally, we assume that births generate an equiprobable Langerin or Rab-11 particle, with a diffusion regime respecting predetermined proportions. The Rab-11 type particles are then generated uniformly in Λ . For the Langerin type particles, we assume that given the pre-jump configuration $\mathbf{x} \in E$, they are generated according

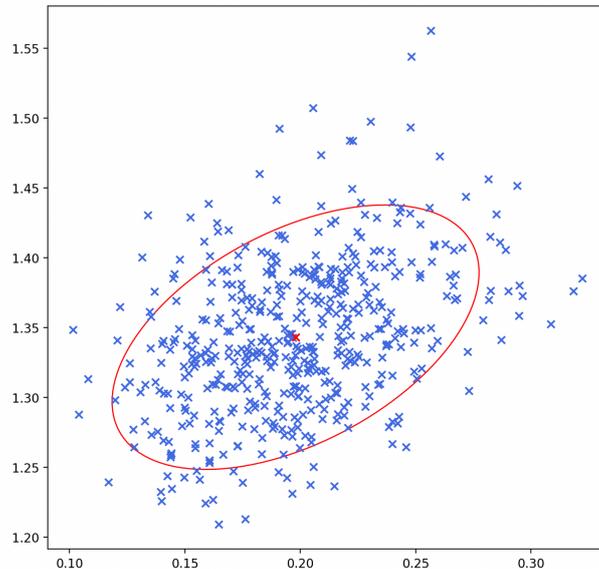


FIGURE 4.2 – Estimation results for $\vartheta = (p, \exp \sigma) = (0.2, 1.34)$ by MLE, based on 500 simulations of the BDMM process specified in Section 4.1. In red : 95% Gaussian confidence ellipsoid, based on the bivariate Gaussian distribution centred at the mean of all estimations, and with covariance matrix the mean of all estimates of the asymptotic covariance matrix $J^{-1}(\vartheta)$.

to the density $k_{\beta}^L(\mathbf{x}, \cdot)$ defined for all $z \in \Lambda$ by

$$k_{\beta}^L(\mathbf{x}, z) = \frac{p}{n_R(\mathbf{x})} \sum_{i=1}^{n_R(\mathbf{x})} \frac{1}{2\pi\sigma^2} \exp\left(\frac{-\|z - z_i^R\|^2}{2\sigma^2}\right) + \frac{(1-p)}{|\Lambda|} \mathbf{1}_{\Lambda}(z), \quad (4.1)$$

where $n_R(\mathbf{x})$ denotes the number of existing Rab-11 particles and z_i^R their spatial coordinates. This density is parametrized by $p \in [0, 1]$ and $\sigma > 0$. It is a convex combination between a mixture of normal distributions centered around the existing Rab-11 particles, with variance σ^2 , like in Example 1 of Chapter 3, and a uniform distribution over Λ . The closer the parameter p is to 1, the more likely Langerin and Rab-11 particles will tend to be colocalized, and the smaller the standard deviation σ is, the stronger the proximity between the two types of particles will be. An illustration of this density is shown in Figure 4.1, for a given configuration of Rab-11 particles represented by red dots.

In line with the application of the next section, we are not interested in the estimation of all features of the process, but only of the parameter $\vartheta = (p, \exp \sigma)$ from

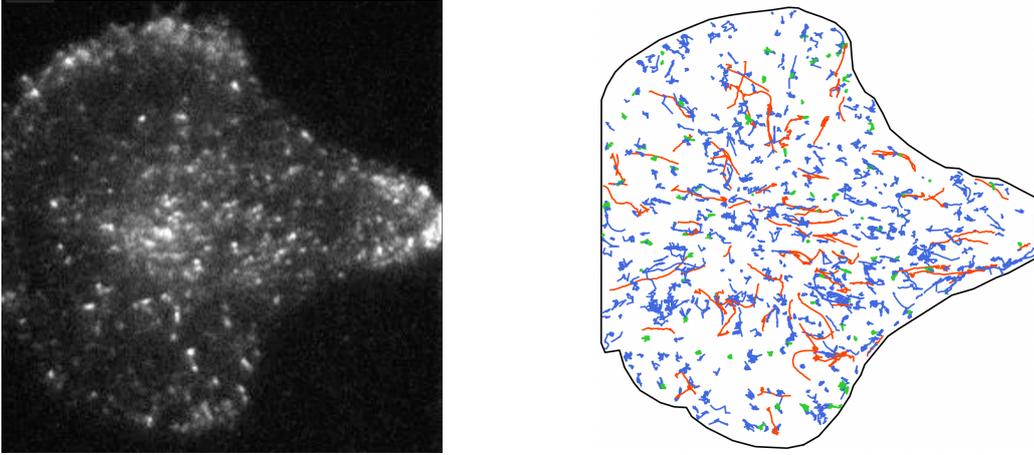


FIGURE 4.3 – Left : a frame from the raw video sequence analysed in Section 4.2, showing Langerin proteins. Right : superposition of all trajectories of Langerin proteins obtained after post-processing (i.e. segmentation and tracking) of the whole raw video. The colors indicate the estimated motion regime (Brownian in blue, directed motion in red and confined motion in green).

the continuous time observation of a single trajectory of the process. We performed 500 simulations of trajectories for $p = 0.2$ and $\exp \sigma = 1.34$, on a time interval similar to the data studied in Section 4.2. Figure 4.2 plots the set of all 500 maximum likelihood estimates of ϑ obtained for the 500 simulations, obtained from the likelihood established in Theorem 1 of Chapter 3. In red, the mean of the 95% confidence ellipsoid is plotted. This is the Gaussian confidence ellipsoid obtained from the bivariate Gaussian distribution centred at the mean of the estimates, and with covariance the mean of the 500 covariance matrices $J_N^{-1}(\hat{\vartheta}_N)$, obtained as explained at the end of Section 3.2.2 of Chapter 3. One can observe that most of the estimates provide a consistent value of the parameter. Also, the mean estimated confidence ellipsoid encodes the estimation uncertainty fairly well.

4.2 Data analysis

The dataset we consider (which are the same as in Chapter 2) correspond to the acquisition of videos with TIRF (Total Internal Reflection Fluorescence) microscopy, depicting two types of proteins observed simultaneously near the membrane of a living cell [BGM⁺14] : Langerin proteins and Rab-11 proteins. One frame of the video of the Langerin proteins only is shown in Figure 4.3 (left). After post-processing

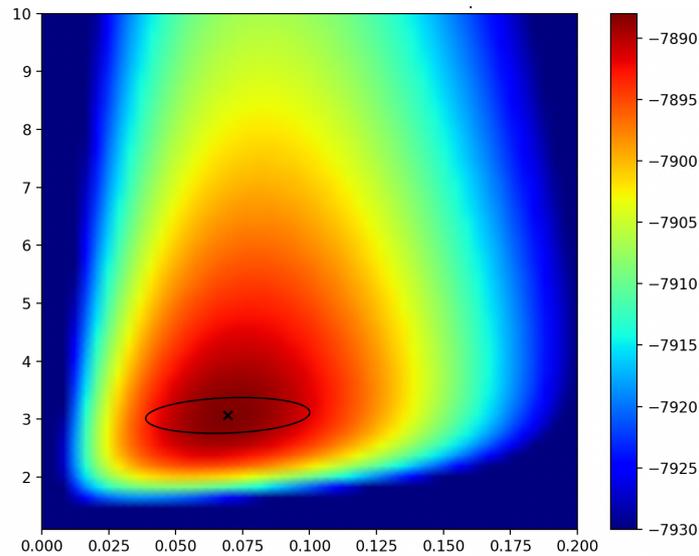


FIGURE 4.4 – Likelihood of the BDMM model for the considered dataset, as a function of the parameters p and $\exp \sigma$. The estimated 95% confidence ellipsoid, centered at the maximum likelihood, is overlaid in black.

following [PBB⁺14], the proteins of interest are detected and represented by a point and tracked by the U-track algorithm [JLM⁺08] along the video sequence to provide trajectories, see Figure 4.3 (right). These trajectories have been further analysed by the method developed in [BKV18] to classify them into three diffusion regimes (these are the colors visible on the Figure 4.3). The full dynamics is in line with a BDMM process; in addition to their displacement, some proteins disappear in the course of time, others appear, and finally some change their diffusion regime, which corresponds in our model to a mutation.

An open biological question is to determine if the Langerin proteins appear close to existing Rab-11 proteins. This colocalization may indicate a strong interaction between these two types of proteins. To address this issue, we estimated by maximum likelihood the birth kernel of the Langerin proteins, assuming that it is of the form (4.1), as in our simulation study. We obtained the estimate $(\hat{p}, \widehat{\exp \sigma}) = (0.07, 3.004)$. The unit of σ is the pixel, the image being of size 250×283 , each pixel representing an area of $80 \times 80 \text{ nm}^2$. Figure 4.4 shows the likelihood value for our dataset, with respect to p and $\log \sigma$, along with the empirical 95% confidence ellipsoid around the maximum, obtained from the estimation of the matrix $J^{-1}(\vartheta)$ explained at the end of Section 3.2.2 in Chapter 3. This estimate suggests that about 7% of

Langerin proteins are colocalized with Rab-11 ones in this example. This proportion is significantly positive in view of the confidence ellipsoid, and each colocalized Langerin protein appears (with high probability) within $3\hat{\sigma} \times 80 = 264$ nm of a Rab-11 protein, where $\hat{\sigma} = \ln(\widehat{\exp \sigma})$.

4.3 Conclusion

In this chapter, we aimed at estimating two parameters of the model : a parameter corresponding to the probability that a new protein will be colocalized with an already existing particle (in the sense that it is born very close to it), and a second parameter characterizing the degree of proximity if the collocation is positive. We first estimated these two parameters on 500 sets of simulations of the same BDMM process, generated using the algorithm described in Part I. We have estimated the 95% confidence ellipse, which was significant and proved the power of the estimator. We then estimated the same two parameters on the real data set already studied in previous chapters. The estimation allowed us to calculate the probability of a Langerin particle being colocalized with a Rab-11 particle at around 7%, with a confidence ellipse that does not contain 0. This confirms colocalization between these two particles, even if it is not very frequent.

Appendix of Part II

REMARK ABOUT THE DIFFUSION COEFFICIENT IN CASE OF CONTINUOUS TIME OBSERVATIONS

If the process $(\mathbf{Z}_t)_{t \geq 0}$ satisfies a stochastic differential equation of the form

$$M^n(\mathbf{z}, \mathbf{m}) : \begin{cases} d\mathbf{Z}_t = \bar{v}_n((\mathbf{Z}_t, \mathbf{m}))dt + \bar{\sigma}_n((\mathbf{Z}_t, \mathbf{m}))dB_t, & t \geq 0, \\ \mathbf{Z}_0 = \mathbf{z} \end{cases}$$

its bracket $\langle \mathbf{Z}, \mathbf{Z} \rangle_t$ satisfies :

$$\int_0^t d\langle \mathbf{Z}, \mathbf{Z} \rangle_s = \int_0^t \bar{\sigma}_n^2(s, (\mathbf{Z}_s, \mathbf{m}))ds.$$

On the other hand, for $0 < t_1^{(p)} < \dots < t_p^{(p)} = t$ a partition of $[0, t]$ with mesh converging to 0, then by approximation of the Riemann integral, we have :

$$\langle \mathbf{Z}, \mathbf{Z} \rangle_t = \lim_{p \rightarrow +\infty} \sum_{k=1}^{p-1} (\mathbf{Z}_{t_{k+1}^{(p)}} - \mathbf{Z}_{t_k^{(p)}})^2$$

which is completely known when the process $(\mathbf{Z}_t)_{0 \leq t \leq T}$ is observed in continuous time.

Therefore the quantity

$$\int_0^t \bar{\sigma}_n^2(s, (\mathbf{Z}_s, \mathbf{m}))ds = \lim_{p \rightarrow +\infty} \sum_{k=1}^p (\mathbf{Z}_{t_{k+1}^{(p)}} - \mathbf{Z}_{t_k^{(p)}})^2$$

can be computed, so its derivative $\left(\int_0^t \bar{\sigma}_n^2(s, (\mathbf{Z}_s, \mathbf{m}))ds \right)' = \bar{\sigma}_n^2(t, (\mathbf{Z}_t, \mathbf{m}))$ is entirely known and ϑ can be deduced. In conclusion it makes no sense to try to estimate ϑ if the process $(\mathbf{Z}_t)_{0 \leq t \leq T}$ is observed in continuous time.

PROOF OF THE LIKELIHOOD EXPRESSION OF A BDMM PROCESS (THEOREM 1 OF CHAPTER 3)

To simplify notation in the proof, we drop the dependence in ϑ of all terms. To get the likelihood, we view $(\mathbf{X}_t)_{t \geq 0}$ as a jump-move process, a more general Markov process than the BDMM process. A jump-move process alternates motions and jumps in E , see [LLGM22]. It is defined through an intensity function α , a jump kernel $K(\mathbf{x}, \cdot)$ and a continuous Markov process \mathbf{Y} on E that drives the motion of the process. For a BDMM, we have $\alpha(\mathbf{x}) = \beta(\mathbf{x}) + \delta(\mathbf{x}) + \tau(\mathbf{x})$ and

$$K(\mathbf{x}, \cdot) = \frac{\beta(\mathbf{x})}{\alpha(\mathbf{x})} K_\beta(\mathbf{x}, \cdot) + \frac{\delta(\mathbf{x})}{\alpha(\mathbf{x})} K_\delta(\mathbf{x}, \cdot) + \frac{\tau(\mathbf{x})}{\alpha(\mathbf{x})} K_\tau(\mathbf{x}, \cdot). \quad (\text{C.1})$$

By our assumptions

$$K(\mathbf{x}, d\mathbf{y}) = k(\mathbf{x}, \mathbf{y}) \nu(\mathbf{x}, d\mathbf{y}), \quad (\text{C.2})$$

where for any $F \in \mathcal{E}$,

$$\nu(\mathbf{x}, F) = \begin{cases} \nu_\beta(\mathbf{x}, F) & \text{if there exist } A \subset \Lambda \text{ and } \mathcal{I} \subset \mathbb{M} \text{ such that } F = \mathbf{x} \cup (A \times \mathcal{I}) \\ \nu_\delta(\mathbf{x}, F) & \text{if } \mathbf{x} \setminus x_i \in F \text{ for some } i \\ \nu_\tau(\mathbf{x}, F) & \text{if there exist } m \in \mathbb{M} \text{ and } i \text{ such that } (\mathbf{x} \setminus (z_i, m_i)) \cup (z_i, m) \in F. \end{cases}$$

In turn,

$$k(\mathbf{x}, \mathbf{y}) = \begin{cases} k_\beta(\mathbf{x}, \mathbf{y}) \frac{\beta(\mathbf{x})}{\alpha(\mathbf{x})} & \text{if there exists } (z, m) \text{ such that } \mathbf{y} = \mathbf{x} \cup (z, m) \\ k_\delta(\mathbf{x}, \mathbf{y}) \frac{\delta(\mathbf{x})}{\alpha(\mathbf{x})} & \text{if } \mathbf{y} = \mathbf{x} \setminus x_i \text{ for some } i \\ k_\tau(\mathbf{x}, \mathbf{y}) \frac{\tau(\mathbf{x})}{\alpha(\mathbf{x})} & \text{if there exists } m \in \mathbb{M} \text{ and } i \text{ such that } \mathbf{y} = (\mathbf{x} \setminus (z_i, m_i)) \cup (z_i, m). \end{cases}$$

Note that a jump-move process observed continuously on $[0, t]$ has by definition a cadlag trajectory on E and can equivalently be described by the vector \underline{X} defined by :

$$\underline{X} = \sum_{n \in \mathbb{N}} (\mathbf{X}_0, \mathbf{X}_{]0, T_1[}, T_1, \mathbf{X}_{T_1}, \mathbf{X}_{]T_1, T_2[}, \dots, T_n - T_{n-1}, \mathbf{X}_{T_n}, \mathbf{X}_{]T_n, t[}) \mathbf{1}_{\{N_t=n\}}. \quad (\text{C.3})$$

We set for any $i \geq 1$, $S_i = T_i - T_{i-1}$, $\mathbf{V}_i = \mathbf{X}_{T_i}$, and we recall that $(\mathbf{X}_t)_{T_i \leq t < T_{i+1}} = (\mathbf{Y}_t^{(i)})_{0 \leq t \leq S_{i+1}}$, where $\mathbf{Y}^{(i)}$ is a continuous Markov process on E , identically distributed as \mathbf{Y} . With these notations \underline{X} is equivalent to \tilde{X} with :

$$\tilde{X} = \sum_{n \in \mathbb{N}} (\mathbf{V}_0, \mathbf{Y}_{[0, S_1]}^{(0)}, S_1, \mathbf{V}_1, \mathbf{Y}_{[0, S_2]}^{(1)}, \dots, S_n, \mathbf{V}_n, \mathbf{Y}_{[0, t - \sum_{j=1}^n S_j]}^{(n)}) \mathbf{1}_{\{N_t=n\}}. \quad (\text{C.4})$$

Note that with this formalism, $\mathbf{Y}_0^{(i)} = \mathbf{V}_i$ for all i , and \tilde{X} takes values in the space $\bigcup_{n \geq 0} \mathcal{W}_n$ where

$$\mathcal{W}_n = \bigotimes_{i=0}^n (E \times \mathcal{C}_{[S_i]} \times \mathbb{R}_+) \times E \times \mathcal{C}_{[t - \sum_{i=0}^n S_i]},$$

and where $\mathcal{C}_{[s]}$ denotes the space of continuous functions from $[0, s]$ to E .

As detailed in Section 3.1.3.3, the continuous Markov process $\mathbf{Y}_s^{|n} = \Pi_n((\mathbf{Z}_s, \mathbf{m}))$ defined on E_n and observed on $[0, t]$, given that $\mathbf{Z}_0 = \mathbf{z}$, admits the density $L(\mathbf{Z}_{[0,t]}, \mathbf{m})$ with respect to the reference process $\Pi_n((\mathbf{U}_t, \mathbf{m}))$ where $\mathbf{U}_t = \mathbf{z} + \int_0^t \bar{\sigma}_n(\mathbf{U}_s, \mathbf{m}) dB_s$. Let us denote by $P_{[t]}^{\mathbf{z}}$ the distribution of this reference process observed on $[0, t]$. For any measurable and bounded function h defined on E_n , we thus have $\mathbb{E}(h(\mathbf{Y}_{[0,t]}^{|n})) = \int_{\mathcal{C}_{[t]}} h(\mathbf{y}) L(\mathbf{y}_{[t]}) dP_{[t]}^{\mathbf{Y}_0^{|n}}(\mathbf{y})$, where $L(\mathbf{y}_{[t]})$ is a short notation for $L(\mathbf{z}_{[0,t]}, \mathbf{m})$ and where $\mathbf{y}_t = (\mathbf{z}_t, \mathbf{m})$ for any $t \geq 0$. Similarly, for $(\mathbf{Y}_t)_{t \geq 0}$ on E defined by $\mathbf{Y}_t = \sum_{n \geq 0} \mathbf{Y}_t^{|n} \mathbf{1}_{\{\mathbf{Y}_0 \in E_n\}}$,

we will write without ambiguity

$$\mathbb{E}(h(\mathbf{Y}_{[0,t]})) = \int_{\mathcal{C}_{[t]}} h(\mathbf{y}) L(\mathbf{y}_{[t]}) dP_{[t]}^{\mathbf{Y}_0}(\mathbf{y}),$$

where the cardinality n defining $P_{[t]}^{\mathbf{Y}_0}$ is implicitly given by $n = n(\mathbf{Y}_0)$.

Let h be a measurable and bounded function. For $n \in \mathbb{N}$, we use the conditional law of S_{n+1} given $(\mathbf{X}_t)_{t \leq T_n}$, T_n and $\mathbf{Y}^{(n)}$, see (3.2), to compute :

$$\begin{aligned} & \mathbb{E}[h(\tilde{X}) \mathbf{1}_{\{N_t=n\}}] \\ &= \mathbb{E} \left[h(\tilde{X}) \mathbf{1}_{\{S_{n+1} > t - \sum_{j=1}^n S_j\}} \mathbf{1}_{\{\sum_{j=1}^n S_j \leq t\}} \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[h(\tilde{X}) \mathbf{1}_{\{S_{n+1} > t - \sum_{j=1}^n S_j\}} \mathbf{1}_{\{\sum_{j=1}^n S_j \leq t\}} \middle| (\mathbf{X}_t)_{t \leq T_n}, T_n, \mathbf{Y}^{(n)} \right] \right] \\ &= \mathbb{E} \left[h(\tilde{X}) \mathbf{1}_{\{\sum_{j=1}^n S_j \leq t\}} \exp \left(- \int_0^{t - \sum_{j=1}^n S_j} \alpha(\mathbf{Y}_u^{(n)}) du \right) \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[h(\tilde{X}) \mathbf{1}_{\{\sum_{j=1}^n S_j \leq t\}} \exp \left(- \int_0^{t - \sum_{j=1}^n S_j} \alpha(\mathbf{Y}_u^{(n)}) du \right) \middle| (\mathbf{X}_t)_{t \leq T_n}, T_n \right] \right]. \end{aligned}$$

Since given S_1, \dots, S_n and $\mathbf{Y}_0^{(n)} = \mathbf{V}_n$, the Markov process $\mathbf{Y}^{(n)}$ observed on the time interval $[0, t - \sum_{j=1}^n S_j]$ admits the density $L(\mathbf{y}_{[t - \sum_{j=1}^n S_j]})$ with respect to $P_{[t - \sum_{j=1}^n S_j]}^{\mathbf{V}_n}$, we have

$$\begin{aligned} & \mathbb{E}[h(\tilde{X}) \mathbf{1}_{\{N_t=n\}}] \\ &= \mathbb{E} \left[\int_{\mathcal{C}_{[t - \sum_{i=1}^n S_j]}} h(\mathbf{V}_0, \mathbf{Y}_{[0, S_1]}^{(0)}, \dots, S_n, \mathbf{V}_n, \mathbf{y}_{[t - \sum_{i=1}^n S_j]}^{(n)}) \mathbf{1}_{\{\sum_{j=1}^n S_j \leq t\}} \right. \\ & \quad \left. \times \exp \left(- \int_0^{t - \sum_{j=1}^n S_j} \alpha(\mathbf{y}_u^{(n)}) du \right) L \left(\mathbf{y}_{[t - \sum_{i=1}^n S_j]}^{(n)} \right) dP_{[t - \sum_{i=1}^n S_j]}^{\mathbf{V}_n}(\mathbf{y}^{(n)}) \right]. \end{aligned}$$

Now, if we condition on $(\mathbf{X}_t)_{t \leq T_{n-1}}$, T_n and $\mathbf{Y}^{(n-1)}$ we know that the post-jump location \mathbf{V}_n admits the density $k(\mathbf{X}_{T_n^-}, \mathbf{v}) = k(\mathbf{Y}_{S_n}^{(n-1)}, \mathbf{v})$ with respect to the measure

$\nu(\mathbf{X}_{T_n^-, \cdot}) = \nu(\mathbf{Y}_{S_n}^{(n-1)}, \cdot)$ on E . Hence

$$\begin{aligned} & \mathbb{E}[h(\tilde{X})\mathbb{1}_{\{N_t=n\}}] \\ &= \mathbb{E} \left[\int_E \int_{\mathcal{C}_{[t-\sum_{i=1}^n s_j]}} h(\mathbf{V}_0, \mathbf{Y}_{[0, S_1[}, \dots, S_n, \mathbf{v}_n, \mathbf{Y}_{[t-\sum_{i=1}^n s_j]}) \mathbb{1}_{\{\sum_{j=1}^n s_j \leq t\}} \right. \\ & \quad \times \exp \left(- \int_0^{t-\sum_{j=1}^n s_j} \alpha(\mathbf{y}_u^{(n)}) du \right) L \left(\mathbf{y}_{[t-\sum_{i=1}^n s_j]}^{(n)} \right) dP_{[t-\sum_{i=1}^n s_j]}^{\mathbf{v}_n}(\mathbf{y}^{(n)}) \\ & \quad \left. \times k(\mathbf{Y}_{S_n}^{(n-1)}, \mathbf{v}_n) \nu(\mathbf{Y}_{S_n}^{(n-1)}, d\mathbf{v}_n) \right]. \end{aligned}$$

Conditioning on $(\mathbf{X}_t)_{t \leq T_{n-1}}$, T_{n-1} and $\mathbf{Y}^{(n-1)}$, we use again the law of S_n given by (3.2) to get

$$\begin{aligned} & \mathbb{E}[h(\tilde{X})\mathbb{1}_{\{N_t=n\}}] \\ &= \mathbb{E} \left[\int_{\mathbb{R}_+} \int_E \int_{\mathcal{C}_{[t-\sum_{i=1}^{n-1} s_j - s_n]}} h(\mathbf{V}_0, \mathbf{Y}_{[0, S_1[}, \dots, s_n, \mathbf{v}_n, \mathbf{Y}_{[t-\sum_{i=1}^{n-1} s_j - s_n]}) \mathbb{1}_{\{\sum_{j=1}^{n-1} s_j + s_n \leq t\}} \right. \\ & \quad \times \exp \left(- \int_0^{t-\sum_{j=1}^{n-1} s_j + s_n} \alpha(\mathbf{y}_u^{(n)}) du \right) L \left(\mathbf{y}_{[t-\sum_{i=1}^{n-1} s_j - s_n]}^{(n)} \right) dP_{[t-\sum_{i=1}^{n-1} s_j - s_n]}^{\mathbf{v}_n}(\mathbf{y}^{(n)}) \\ & \quad \left. \times k(\mathbf{Y}_{s_n}^{(n-1)}, \mathbf{v}_n) \nu(\mathbf{Y}_{s_n}^{(n-1)}, d\mathbf{v}_n) \alpha(\mathbf{Y}_{s_n}^{(n-1)}) \exp \left(- \int_0^{s_n} \alpha(\mathbf{Y}_u^{(n-1)}) du \right) ds_n \right]. \end{aligned}$$

We can continue successively this process, using the (conditional) density of each move process $\mathbf{Y}^{(i)}$ on $[0, S_i]$, of each post-jump location \mathbf{V}_i , and of each inter-jump time S_i , and we obtain

$$\begin{aligned} & \mathbb{E}[h(\tilde{X})\mathbb{1}_{\{N_t=n\}}] \\ &= \int_{\mathcal{W}_n} h(\mathbf{v}_0, \mathbf{y}_{[s_1]}^{(0)}, \dots, s_n, \mathbf{v}_n, \mathbf{Y}_{[t-\sum_{i=1}^n s_j]}^{(n)}) \mathbb{1}_{\{\sum_{j=1}^n s_j \leq t\}} \\ & \quad \times \exp \left(- \int_0^{t-\sum_{j=1}^n s_j} \alpha(\mathbf{y}_u^{(n)}) du \right) \times L(\mathbf{y}_{[t-\sum_{i=1}^n s_j]}^{(n)}) dP_{[t-\sum_{i=1}^n s_j]}^{\mathbf{v}_n}(\mathbf{y}^{(n)}) \\ & \quad \times k(\mathbf{y}_{s_n}^{(n-1)}, \mathbf{v}_n) \nu(\mathbf{y}_{s_n}^{(n-1)}, d\mathbf{v}_n) \alpha(\mathbf{y}_{s_n}^{(n-1)}) \exp \left(- \int_0^{s_n} \alpha(\mathbf{y}_u^{(n-1)}) du \right) ds_n \\ & \quad \times \prod_{i=1}^{n-1} \left[L(\mathbf{y}_{[s_{i+1}]}^{(i)}) dP_{[s_{i+1}]}^{\mathbf{v}_i}(\mathbf{y}^{(i)}) k(\mathbf{y}_{s_i}^{(i-1)}, \mathbf{v}_i) \nu(\mathbf{y}_{s_i}^{(i-1)}, d\mathbf{v}_i) \alpha(\mathbf{y}_{s_i}^{(i-1)}) \right. \\ & \quad \left. \times \exp \left(- \int_0^{s_i} \alpha(\mathbf{y}_u^{(i-1)}) du \right) ds_i \right] L(\mathbf{y}_{[s_1]}^{(0)}) dP_{[s_1]}^{\mathbf{v}_0}(\mathbf{y}^{(0)}) p(d\mathbf{v}_0), \end{aligned}$$

where p denotes the distribution of the initial state on E . If we come back to the initial expression (C.3) of \underline{X} from that of \tilde{X} , using the shortcut $L(\mathbf{x}_{[t_i, t_{i+1}[})$ for $L(\mathbf{y}_{[s_{i+1}]}) = L(\mathbf{z}_{[0, t_{i+1}-t_i], \mathbf{m}^{(i)})$ in view of $\mathbf{x}_t = \{(\mathbf{z}_{t-T_i}^{(i)}, \mathbf{m}^{(i)})\}$ for all $t \in [t_i, t_{i+1}[$, we obtain

$$\begin{aligned} \mathbb{E}[h(\underline{X})\mathbb{1}_{\{N_t=n\}}] &= \int_{\mathcal{W}_n} h(\mathbf{x}_0, \mathbf{x}_{[0, t_1[}, \dots, t_n - t_{n-1}, \mathbf{x}_{t_n}, \mathbf{x}_{[t_n, t]}) \mathbb{1}_{t_1 \leq \dots \leq t_n \leq t} \\ &\times \exp\left(-\int_0^t \alpha(\mathbf{x}_u) du\right) L(\mathbf{x}_{[t_n, t]}) \prod_{i=1}^n \left(L(\mathbf{x}_{[t_{i-1}, t_i[}) k(\mathbf{x}_{t_i^-}, \mathbf{x}_{t_i}) \alpha(\mathbf{x}_{t_i^-})\right) \eta^{(n)}(d\underline{x}), \end{aligned}$$

where $\eta^{(n)}$ is the measure given for $\underline{x} = (\mathbf{x}_0, \mathbf{x}_{[0, t_1[}, \dots, t_n - t_{n-1}, \mathbf{x}_{t_n}, \mathbf{x}_{[t_n, t]})$ by

$$\begin{aligned} \eta^{(n)}(d\underline{x}) &= \mathbb{1}_{t_1 \leq \dots \leq t_n \leq t} dP_{[t_n, t]}^{\mathbf{x}_{t_n}}(\mathbf{x}) \nu(\mathbf{x}_{t_n^-}, d\mathbf{x}_{t_n}) dt_n \\ &\times \prod_{i=1}^{n-1} \left(dP_{[t_i, t_{i+1}[}^{\mathbf{x}_{t_i}}(\mathbf{x}) \nu(\mathbf{x}_{t_i^-}, \mathbf{x}_{t_i}) dt_i\right) dP_{[0, t_1[}^{\mathbf{x}_0}(\mathbf{x}) p(d\mathbf{x}_0). \end{aligned}$$

Here, just as we have used the natural notation $L(\mathbf{x}_{[t_i, t_{i+1}[})$ for $L(\mathbf{y}_{[s_{i+1}]})$, we use $dP_{[t_i, t_{i+1}[}^{\mathbf{x}_{t_i}}(\mathbf{x})$ for $dP_{[s_{i+1}]})^{\mathbf{y}^{(i)}}(\mathbf{y}^{(i)})$.

Since for any h , $\mathbb{E}(h(\underline{X})) = \sum_{n \geq 0} \mathbb{E}(h(\underline{X})\mathbb{1}_{\{N_t=n\}})$, we have proven that the likelihood of \underline{X} given by (C.3), with respect to the underlying measure $\eta^{(\cdot)} = \sum_{n \in \mathbb{N}} \eta^{(n)}(\cdot \cap \mathcal{W}_n)$ is given by

$$\mathcal{L}(\underline{X}) = \exp\left(-\int_0^t \alpha(\mathbf{X}_u) du\right) L(\mathbf{X}_{[t, N_t, t]}) \prod_{i=1}^{N_t} \left(L(\mathbf{X}_{[T_{i-1}, T_i[}) k(\mathbf{X}_{T_i^-}, \mathbf{X}_{T_i}) \alpha(\mathbf{X}_{T_i^-})\right).$$

If we replace α and k by their specific expression in (C.2), we obtain the result in the case of a BDMM process. Given that there is a one to one correspondance between the representation (C.3) and the trajectory $(\mathbf{X}_s)_{0 \leq s \leq t}$, we can view the above likelihood as the likelihood of $(\mathbf{X}_s)_{0 \leq s \leq t}$, where η translates to a measure on the space of cadlag function from $[0, t]$ to E .

PROOF OF THE LAN PROPERTY OF A BDMM PROCESS (THEOREM 2 OF CHAPTER 3)

Let us fix $N \geq 1$ and $h \in \mathbb{R}^\ell$. We let $\vartheta_N = \vartheta + h/\sqrt{N}$ and we denote

$$\Lambda^{\vartheta_N/\vartheta}(t) = \log \frac{\mathcal{L}_t(\vartheta + \frac{h}{\sqrt{N}})}{\mathcal{L}_t(\vartheta)}.$$

By Theorem 1 the log-likelihood ratio process is given by :

$$\begin{aligned} \Lambda^{\vartheta_N/\vartheta}(t) = & \Lambda_L^{\vartheta_N/\vartheta}(t) + \Lambda_\beta^{\vartheta_N/\vartheta}(t) + \Lambda_\delta^{\vartheta_N/\vartheta}(t) + \Lambda_\tau^{\vartheta_N/\vartheta}(t) \\ & + \Lambda_{k_\beta}^{\vartheta_N/\vartheta}(t) + \Lambda_{k_\delta}^{\vartheta_N/\vartheta}(t) + \Lambda_{k_\tau}^{\vartheta_N/\vartheta}(t), \end{aligned} \quad (\text{D.1})$$

where

$$\Lambda_\gamma^{\vartheta_N/\vartheta}(t) = \int_0^t \log \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} \right) dN_s^\gamma - \int_0^t (\gamma(\vartheta_N, \mathbf{X}_s) - \gamma(\vartheta, \mathbf{X}_s)) ds, \quad (\text{D.2})$$

$$\Lambda_{k_\gamma}^{\vartheta_N/\vartheta}(t) = \int_0^t \log \left(\frac{k_\gamma(\vartheta_N, \mathbf{X}_{s^-}, \mathbf{X}_s)}{k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)} \right) dN_s^\gamma, \quad (\text{D.3})$$

$$\Lambda_L^{\vartheta_N/\vartheta}(t) = \log \frac{L(\vartheta_N, \mathbf{X}_{[T_{N_t}, t]})}{L(\vartheta, \mathbf{X}_{[T_{N_t}, t]})} + \sum_{i=0}^{N_t-1} \log \frac{L(\vartheta_N, \mathbf{X}_{[T_i, T_{i+1}])}}{L(\vartheta, \mathbf{X}_{[T_i, T_{i+1}])}. \quad (\text{D.4})$$

The proof of Theorem 2 follows a standard scheme as in [Lus92] and [Löc02a]. The main step consists in proving the LAN decomposition for each of the three terms (D.2), (D.3) and (D.4) above, which is the statement of the three following lemmas. Then (3.17) is deduced by combining these decompositions, as carried out next.

Lemma 1. *Under the same assumptions as in Theorem 2, we have for any $t > 0$, $h \in \mathbb{R}^\ell$ and $\vartheta \in \Theta$,*

$$\Lambda_\gamma^{\vartheta N/\vartheta}(tN) = h^T M_\gamma^{N,\vartheta}(t) - \frac{1}{2} h^T \langle M_\gamma^{N,\vartheta} \rangle(t) h + \text{Rem}_{N,\vartheta,h}^\gamma(t), \quad (\text{D.5})$$

where $\text{Rem}_{N,\vartheta,h}^\gamma(t)$ tends to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability when $N \rightarrow +\infty$, and $(M_\gamma^{N,\vartheta}(t))_{t>0}$ is a martingale with respect to \mathcal{F}_{tN} given by (3.9).

Lemma 2. *Under the same assumptions as in Theorem 2, we have for any $t > 0$, $h \in \mathbb{R}^\ell$ and $\vartheta \in \Theta$,*

$$\Lambda_{k_\gamma}^{\vartheta N/\vartheta}(tN) = h^T M_{k_\gamma}^{N,\vartheta}(t) - \frac{1}{2} h^T \langle M_{k_\gamma}^{N,\vartheta} \rangle(t) h + \text{Rem}_{N,\vartheta,h}^{k_\gamma}(t), \quad (\text{D.6})$$

where $\text{Rem}_{N,\vartheta,h}^{k_\gamma}(t)$ tends to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability when $N \rightarrow +\infty$, and $(M_{k_\gamma}^{N,\vartheta}(t))_{t>0}$ is a martingale with respect to \mathcal{F}_{tN} given by (3.10).

Lemma 3. *Under the same assumptions as in Theorem 2, we have for any $t > 0$, $h \in \mathbb{R}^\ell$ and $\vartheta \in \Theta$,*

$$\Lambda_L^{\vartheta N/\vartheta}(tN) = h^T M_L^{N,\vartheta}(t) - \frac{1}{2} h^T \langle M_L^{N,\vartheta} \rangle(t) h + \text{Rem}_{N,\vartheta,h}^L(t), \quad (\text{D.7})$$

where $\text{Rem}_{N,\vartheta,h}^L(t)$ tends to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability when $N \rightarrow +\infty$, and $(M_L^{N,\vartheta}(t))_{t>0}$ is a martingale with respect to \mathcal{F}_{tN} given by (3.11).

The proofs of these three lemmas are postponed to the end of this section. Let us deduce (3.17). From their statements and (D.1), we have

$$\Lambda^{\vartheta N/\vartheta}(tN) = h^T M^{N,\vartheta}(t) - \frac{1}{2} h^T \sum_{i \in \mathcal{S}} \langle M_i^{N,\vartheta} \rangle(t) h + \text{Rem}_{N,\vartheta,h}(t),$$

where $\mathcal{S} = \{L, \beta, \delta, \tau, k_\beta, k_\delta, k_\tau\}$, $M^{N,\vartheta}(t) = \sum_{i \in \mathcal{S}} M_i^{N,\vartheta}$ and $\text{Rem}_{N,\vartheta,h}(t) = \sum_{i \in \mathcal{S}} \text{Rem}_{N,\vartheta,h}^i(t)$. By Lemmas 1, 2 and 3, we know that $\text{Rem}_{N,\vartheta,h}(t)$ tends to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability. To show the decomposition (3.17), it remains to prove that

$$\langle M^{N,\vartheta} \rangle(t) = \sum_{i \in \mathcal{S}} \langle M_i^{N,\vartheta} \rangle(t),$$

which, given the definition of $M^{N,\vartheta}(t)$, boils down to proving that

$$\forall i, j \in \mathcal{S}, i \neq j, \quad \langle M_i^{N,\vartheta}, M_j^{N,\vartheta} \rangle(t) = 0. \quad (\text{D.8})$$

First, for $i = L$, since $M_L^{N,\vartheta}(t)$ is a continuous martingale with finite variations, we have that $\langle M_L^{N,\vartheta}, M_j^{N,\vartheta} \rangle(t) = 0$ for any $j \neq L$, see [Kle12]. Second, for all $i, j \in \mathcal{S} \setminus \{L\}$, $i \neq j$, we have

$$[M_i^{N,\vartheta}, M_j^{N,\vartheta}](t) = \sum_{s \leq t} \Delta M_i^{N,\vartheta}(s) \Delta M_j^{N,\vartheta}(s),$$

where $\Delta M_i^{N,\vartheta}(s) = M_i^{N,\vartheta}(s) - M_i^{N,\vartheta}(s^-)$. For $\gamma \neq \gamma'$, $M_\gamma^{N,\vartheta}(t)$ and $M_{\gamma'}^{N,\vartheta}(t)$ do not have any jump in common, and similarly for $M_\gamma^{N,\vartheta}(t)$ and $M_{k_{\gamma'}}^{N,\vartheta}(t)$, since a birth, a death or a mutation cannot occur at the same moment almost surely. Consequently $[M_\gamma^{N,\vartheta}, M_{\gamma'}^{N,\vartheta}](t) = [M_\gamma^{N,\vartheta}, M_{k_{\gamma'}}^{N,\vartheta}](t) = 0$ for any $\gamma \neq \gamma'$. This implies (D.8) for $i = \gamma$ and $j \in \{\gamma', k_{\gamma'}\}$, $\gamma \neq \gamma'$. To complete the proof of (D.8), it remains to show that $\langle M_\gamma^{N,\vartheta}, M_{k_\gamma}^{N,\vartheta} \rangle(t) = 0$. We have that

$$[M_\gamma^{N,\vartheta}, M_{k_\gamma}^{N,\vartheta}](t) = \frac{1}{N} \int_0^{tN} \frac{\nabla_1 \gamma(\vartheta, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} \frac{(\nabla_1 k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s))^T}{k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)} dN_s^\gamma,$$

so that by Lemma 4, the associated angle bracket (corresponding to the compensator of the square bracket, see [Kle12]) is given by

$$\langle M_\gamma^{N,\vartheta}, M_{k_\gamma}^{N,\vartheta} \rangle(t) = \frac{1}{N} \int_0^{tN} \nabla_1 \gamma(\vartheta, \mathbf{X}_s) \int_E (\nabla_1 k_\gamma(\vartheta, \mathbf{X}_s, \mathbf{y}))^T \nu_\gamma(\mathbf{X}_s, d\mathbf{y}) ds.$$

Remember that for any $\vartheta \in \Theta$ and any $\mathbf{x} \in E$, $\int_E k_\gamma(\vartheta, \mathbf{x}, \mathbf{y}) \nu(\mathbf{x}, d\mathbf{y}) = 1$. This implies $\int_E \nabla_1 k_\gamma(\vartheta, \mathbf{x}, \mathbf{y}) \nu_\gamma(\mathbf{x}, d\mathbf{y}) = 0$ if interchange of integration and differentiation is valid. In particular, it can be showed that the latter holds under Condition **B**(k_γ, γ). Consequently, $\langle M_\gamma^{N,\vartheta}, M_{k_\gamma}^{N,\vartheta} \rangle(t) = 0$ and (D.8) is proven, which completes the proof of (3.17).

The last statement of Theorem 2, that is the convergence (3.18), is a direct consequence of Theorems 4.12 and 4.22 of [HL00], thanks to our hypothesis **(H)** that implies that \emptyset is a recurrent atom for $(\mathbf{X}_t)_{t>0}$.

D.1 Proof of Lemma 1

Let us prove that $Rem_{N,\vartheta,h}^\gamma(t)$ converges to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability, where

$$Rem_{N,\vartheta,h}^\gamma(t) = \Lambda_\gamma^{\vartheta_N/\vartheta}(tN) - h^T M_\gamma^{N,\vartheta}(t) + \frac{1}{2} h^T \langle M_\gamma^{N,\vartheta} \rangle(t) h,$$

and $M_\gamma^{N,\vartheta}(t)$ is given by (3.9). From (D.2), we get

$$\begin{aligned} \Lambda_\gamma^{\vartheta_N/\vartheta}(tN) &= \int_0^{tN} \left[\log \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} \right) - \frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} + 1 \right] dN_s^\gamma \\ &\quad + \int_0^{tN} \left[\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 \right] dM_s^\gamma, \end{aligned}$$

where $M_t^\gamma = N_t^\gamma - \int_0^t \gamma(\vartheta, \mathbf{X}_s) ds$ is a \mathcal{F}_t -martingale. We have in particular $[M^\gamma]_t = N_t$ and $\langle M^\gamma \rangle_t = \int_0^t \gamma(\vartheta, \mathbf{X}_s) ds$, see for instance [Kle12]. We deduce that

$$\begin{aligned} [M_\gamma^{N,\vartheta}](t) &= \frac{1}{N} \int_0^{tN} \frac{(\nabla_1 \gamma)(\nabla_1 \gamma)^T}{\gamma^2}(\vartheta, \mathbf{X}_{s^-}) d[M^\gamma]_s \\ &= \frac{1}{N} \int_0^{tN} \frac{(\nabla_1 \gamma)(\nabla_1 \gamma)^T}{\gamma^2}(\vartheta, \mathbf{X}_{s^-}) dN_s^\gamma. \end{aligned}$$

Note that $\gamma(\vartheta, \mathbf{X}_{T_i^-}) > 0$ for all γ -jump time T_i , so that the previous formula remains true with the addition of $\mathbf{1}_{\gamma(\vartheta, \mathbf{X}_{s^-}) > 0}$ in the integrand. So we get

$$\langle M_\gamma^{N,\vartheta} \rangle(t) = \frac{1}{N} \int_0^{tN} \mathbf{1}_{\gamma > 0} \frac{(\nabla_1 \gamma)(\nabla_1 \gamma)^T}{\gamma}(\vartheta, \mathbf{X}_s) ds,$$

as claimed in (3.12), which well satisfies the property that $[M_\gamma^{N,\vartheta}](t) - \langle M_\gamma^{N,\vartheta} \rangle(t)$ is a martingale. Using these representations, we may write

$$\begin{aligned} Rem_{N,\vartheta,h}^\gamma(t) &= \Lambda_\gamma^{\vartheta_N/\vartheta}(tN) - h^T M_\gamma^{N,\vartheta}(t) + \frac{1}{2} h^T \left(\langle M_\gamma^{N,\vartheta} \rangle(t) - [M_\gamma^{N,\vartheta}](t) \right) h + \frac{1}{2} h^T [M_\gamma^{N,\vartheta}](t) h \\ &= S_t^N + A_t^N + R_t^N, \end{aligned}$$

where

$$S_t^N = \int_0^{tN} \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 - \frac{h^T \nabla_1 \gamma(\vartheta, \mathbf{X}_{s^-})}{\sqrt{N} \gamma(\vartheta, \mathbf{X}_{s^-})} \right) dM_s^\gamma,$$

$$\begin{aligned}
 A_t^N &= \frac{1}{2N} \int_0^{tN} \mathbf{1}_{\gamma>0} \frac{(h^T \nabla_1 \gamma)^2}{\gamma}(\vartheta, \mathbf{X}_s) ds - \frac{1}{2N} \int_0^{tN} \frac{(h^T \nabla_1 \gamma)^2}{\gamma^2}(\vartheta, \mathbf{X}_{s-}) dN_s^\gamma, \\
 R_t^N &= \int_0^{tN} \left(\log \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s-})}{\gamma(\vartheta, \mathbf{X}_{s-})} \right) - \frac{\gamma(\vartheta_N, \mathbf{X}_{s-})}{\gamma(\vartheta, \mathbf{X}_{s-})} + 1 + \frac{1}{2N} \frac{(h^T \nabla_1 \gamma)^2}{\gamma^2}(\vartheta, \mathbf{X}_{s-}) \right) dN_s^\gamma.
 \end{aligned}$$

To complete the proof, we show below that each of these three terms tends to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability.

First step : Convergence of S_t^N to 0 in probability.

By the Markov inequality, we have for any $a > 0$,

$$\mathbb{P}_{\mathbf{x}_0}^\vartheta \left(|S_t^N| > a \right) \leq \frac{1}{a^2} \mathbb{E}_{\mathbf{x}_0}^\vartheta \left((S_t^N)^2 \right) = \frac{1}{a^2} \mathbb{E}_{\mathbf{x}_0}^\vartheta \left(\langle S^N \rangle_t \right),$$

where the last equality comes from the fact that $(S_t^N)^2 - \langle S^N \rangle_t$ is by definition a centred martingale (see [Kle12]). We prove in the following that $\mathbb{E}_{\mathbf{x}_0}^\vartheta \left(\langle S^N \rangle_t \right)$ tends to 0. We have

$$\begin{aligned}
 [S^N]_t &= \int_0^{tN} \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s-})}{\gamma(\vartheta, \mathbf{X}_{s-})} - 1 - \frac{h^T}{\sqrt{N}} \frac{\nabla_1 \gamma(\vartheta, \mathbf{X}_{s-})}{\gamma(\vartheta, \mathbf{X}_{s-})} \right)^2 dN_s^\gamma \\
 &= \int_0^{tN} \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s-})}{\gamma(\vartheta, \mathbf{X}_{s-})} - 1 - \frac{h^T}{\sqrt{N}} \frac{\nabla_1 \gamma(\vartheta, \mathbf{X}_{s-})}{\gamma(\vartheta, \mathbf{X}_{s-})} \right)^2 \mathbf{1}_{\gamma(\vartheta, \mathbf{X}_{s-})>0} dN_s^\gamma,
 \end{aligned}$$

so that

$$\langle S^N \rangle_t = \int_0^{tN} \left(\frac{\gamma(\vartheta_N, \mathbf{X}_s)}{\gamma(\vartheta, \mathbf{X}_s)} - 1 - \frac{h^T}{\sqrt{N}} \frac{\nabla_1 \gamma(\vartheta, \mathbf{X}_s)}{\gamma(\vartheta, \mathbf{X}_s)} \right)^2 \mathbf{1}_{\gamma(\vartheta, \mathbf{X}_s)>0} \gamma(\vartheta, \mathbf{X}_s) ds. \quad (\text{D.9})$$

Note that for any $\vartheta, \vartheta' \in \Theta$ and any \mathbf{x} ,

$$\gamma(\vartheta', \mathbf{x}) - \gamma(\vartheta, \mathbf{x}) = \int_0^1 (\nabla_1 \gamma(\vartheta + t(\vartheta' - \vartheta), \mathbf{x}))^T (\vartheta' - \vartheta) dt.$$

Thus we have for all $\mathbf{x} \in E$ and $\vartheta, \vartheta' \in \Theta$ such that $\gamma(\vartheta, \mathbf{x}) > 0$:

$$\begin{aligned}
 &\left(\frac{\gamma(\vartheta', \mathbf{x})}{\gamma(\vartheta, \mathbf{x})} - 1 - \frac{(\nabla_1 \gamma(\vartheta, \mathbf{x}))^T (\vartheta' - \vartheta)}{\gamma(\vartheta, \mathbf{x})} \right)^2 \\
 &= \left(\int_0^1 \frac{(\nabla_1 \gamma(\vartheta + t(\vartheta' - \vartheta), \mathbf{x}) - \nabla_1 \gamma(\vartheta, \mathbf{x}))^T (\vartheta' - \vartheta)}{\gamma(\vartheta, \mathbf{x})} dt \right)^2 \\
 &\leq \|\vartheta' - \vartheta\|^2 f_\vartheta^\gamma(\mathbf{x}, \|\vartheta' - \vartheta\|), \quad (\text{D.10})
 \end{aligned}$$

where $f_\vartheta^\gamma(\mathbf{x}, \rho)$ is given by (3.6). Applying this inequality in (D.9) with $\vartheta' = \vartheta_N = \vartheta + h/\sqrt{N}$, we obtain

$$\langle S^N \rangle_t \leq \frac{\|h\|^2}{N} \int_0^{tN} f_\vartheta^\gamma \left(\mathbf{X}_s, \frac{\|h\|}{\sqrt{N}} \right) \gamma(\vartheta, \mathbf{X}_s) ds. \quad (\text{D.11})$$

Since $f_\vartheta^\gamma(\mathbf{x}, \cdot)$ is increasing, we have by condition **B**(γ) that for any $\vartheta \in \Theta$ and N large enough,

$$\sup_{\mathbf{x} \in E} \left| f_\vartheta^\gamma \left(\mathbf{x}, \frac{\|h\|}{\sqrt{N}} \right) \gamma(\vartheta, \mathbf{x}) \right| < d(\vartheta), \quad (\text{D.12})$$

for some $d(\vartheta) > 0$. We can thus apply the ergodic assumption **(H)** to the right-hand side of (D.11) to get

$$\mathbb{E}_{\mathbf{x}_0}^\vartheta \left(|\langle S^N \rangle_t| \right) \leq \frac{\|h\|^2 c_\vartheta(\mathbf{x}_0) d(\vartheta)}{r_\vartheta N} + \|h\|^2 t \int_E \left| f_\vartheta^\gamma \left(\mathbf{x}, \frac{\|h\|}{\sqrt{N}} \right) \gamma(\vartheta, \mathbf{x}) \right| \mu_\infty^\vartheta(d\mathbf{x}).$$

In view of (D.12), and since for any $\mathbf{x} \in E$, $f_\vartheta^\gamma(\mathbf{x}, \|h\|/\sqrt{N}) \rightarrow 0$ when $N \rightarrow \infty$, we can conclude by the dominated convergence theorem.

Second step : Convergence of A_t^N to 0 in probability.

Let $\epsilon > 0$. We have $2A_t^N = A_1 - A_2 - A_3$ where

$$\begin{aligned} A_1 &= \frac{1}{N} \int_0^{tN} \mathbf{1}_{\gamma > 0} \frac{(h^T \nabla_1 \gamma)^2}{\gamma}(\vartheta, \mathbf{X}_s) \mathbf{1}_{\left\{ \left| \frac{h^T \nabla_1 \gamma}{\gamma}(\vartheta, \mathbf{X}_s) \right| > \epsilon \sqrt{N} \right\}} ds, \\ A_2 &= \frac{1}{N} \int_0^{tN} \frac{(h^T \nabla_1 \gamma)^2}{\gamma^2}(\vartheta, \mathbf{X}_{s^-}) \mathbf{1}_{\left\{ \left| \frac{h^T \nabla_1 \gamma}{\gamma}(\vartheta, \mathbf{X}_{s^-}) \right| > \epsilon \sqrt{N} \right\}} dN_s^\gamma, \\ A_3 &= \frac{1}{N} \int_0^{tN} \mathbf{1}_{\gamma > 0} \frac{(h^T \nabla_1 \gamma)^2}{\gamma^2}(\vartheta, \mathbf{X}_{s^-}) \mathbf{1}_{\left\{ \left| \frac{h^T \nabla_1 \gamma}{\gamma}(\vartheta, \mathbf{X}_{s^-}) \right| \leq \epsilon \sqrt{N} \right\}} dM_s^\gamma. \end{aligned} \quad (\text{D.13})$$

Note that $\mathbb{E}_{\mathbf{x}_0}^\vartheta(|A_1|) = \mathbb{E}_{\mathbf{x}_0}^\vartheta(A_1) = \mathbb{E}_{\mathbf{x}_0}^\vartheta(A_2) = \mathbb{E}_{\mathbf{x}_0}^\vartheta(|A_2|)$. We prove below that for any $\epsilon > 0$, $\mathbb{E}_{\mathbf{x}_0}^\vartheta(|A_1|) \rightarrow 0$ as $N \rightarrow \infty$ and then we deal with the convergence in probability of A_3 .

By condition **A**(γ), we know that given $\vartheta \in \Theta$, the integrand in A_1 is uniformly bounded in \mathbf{X}_s by some $d(\vartheta) > 0$. We can then apply **(H)** to get

$$\mathbb{E}_{\mathbf{x}_0}^\vartheta(|A_1|) \leq \frac{c_\vartheta(\mathbf{x}_0) d(\vartheta)}{r_\vartheta N} + t \int_E \left| \mathbf{1}_{\gamma > 0} \frac{(h^T \nabla_1 \gamma)^2}{\gamma}(\vartheta, \mathbf{x}) \mathbf{1}_{\left\{ \left| \frac{h^T \nabla_1 \gamma}{\gamma}(\vartheta, \mathbf{x}) \right| > \epsilon \sqrt{N} \right\}} \right| \mu_\infty^\vartheta(d\mathbf{x}).$$

This integrand is bounded thanks to condition $\mathbf{A}(\gamma)$ and converges to 0 as $N \rightarrow \infty$ for any given \mathbf{x} and ϑ . We thus conclude by the dominated convergence theorem that $\mathbb{E}_{\mathbf{x}_0}^\vartheta(|A_1|) \rightarrow 0$ as $N \rightarrow \infty$.

Concerning A_3 , since it is a centred martingale, we have for any $a > 0$,

$$\mathbb{P}_{\mathbf{x}_0}^\vartheta(|A_3| > a) \leq \frac{1}{a^2} \mathbb{E}_{\mathbf{x}_0}^\vartheta(|A_3|^2) = \frac{1}{a^2} \mathbb{E}_{\mathbf{x}_0}^\vartheta(\langle A_3 \rangle)$$

where

$$\begin{aligned} \langle A_3 \rangle &= \frac{1}{N^2} \int_0^{tN} \mathbf{1}_{\gamma>0} \frac{(h^T \nabla_1 \gamma)^4}{\gamma^3}(\vartheta, \mathbf{X}_s) \mathbf{1}_{\left\{ \left| \frac{h^T \nabla_1 \gamma}{\gamma}(\vartheta, \mathbf{X}_s) \right| \leq \epsilon \sqrt{N} \right\}} ds \\ &\leq \frac{\epsilon^2}{N} \int_0^{tN} \mathbf{1}_{\gamma>0} \frac{(h^T \nabla_1 \gamma)^2}{\gamma}(\vartheta, \mathbf{X}_s) ds. \end{aligned}$$

We can again apply (\mathbf{H}) to the latest integral, thanks to condition $\mathbf{A}(\gamma)$, and we obtain

$$\mathbb{E}_{\mathbf{x}_0}^\vartheta(\langle A_3 \rangle) \leq \epsilon^2 C(\mathbf{x}_0, \vartheta, h)$$

where $C(\mathbf{x}_0, \vartheta, h) > 0$. So for any $a > 0$ and any $\eta > 0$, we can choose ϵ so that $\mathbb{P}_{\mathbf{x}_0}^\vartheta(|A_3| > a) \leq \eta$. For these choices, we can further choose N large enough so that $\mathbb{P}_{\mathbf{x}_0}^\vartheta(|A_1| > a) \leq \eta$ and $\mathbb{P}_{\mathbf{x}_0}^\vartheta(|A_2| > a) \leq \eta$, because for any $\epsilon > 0$, $\mathbb{E}_{\mathbf{x}_0}^\vartheta(|A_1|) = \mathbb{E}_{\mathbf{x}_0}^\vartheta(|A_2|) \rightarrow 0$ as $N \rightarrow \infty$. This entails the convergence of A_t^N to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability.

Third step : Convergence of R_t^N to 0 in probability.

Denoting by $\varphi : x > -1 \mapsto \log(1+x) - x + x^2/2$, we have $R_t^N = R_1 - R_2/2$ with

$$\begin{aligned} R_1 &= \int_0^{tN} \varphi \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s-})}{\gamma(\vartheta, \mathbf{X}_{s-})} - 1 \right) dN_s^\gamma, \\ R_2 &= \int_0^{tN} \left(\left[\frac{\gamma(\vartheta_N, \mathbf{X}_{s-})}{\gamma(\vartheta, \mathbf{X}_{s-})} - 1 \right]^2 - \left[\frac{h^T \nabla_1 \gamma(\vartheta, \mathbf{X}_{s-})}{\sqrt{N} \gamma(\vartheta, \mathbf{X}_{s-})} \right]^2 \right) dN_s^\gamma. \end{aligned} \quad (\text{D.14})$$

We start by proving that R_2 tends to 0 in probability. Write $R_2 = \int (f^2 - g^2) dN_s^\gamma$ with obvious notations. We have by the Cauchy-Schwartz inequality and the fact that $(f+g)^2 = (f-g+2g)^2 \leq 2(f-g)^2 + 8g^2$,

$$\begin{aligned} R_2^2 &\leq \int (f-g)^2 dN_s^\gamma \int (f+g)^2 dN_s^\gamma \\ &\leq \left(\int (f-g)^2 dN_s^\gamma \right) \left(8 \int (f-g)^2 dN_s^\gamma + 8 \int g^2 dN_s^\gamma \right). \end{aligned}$$

To prove the convergence of R_2 , it is then sufficient to show that $\int (f - g)^2 dN_s^\gamma = o_{\mathbb{P}_{\mathbf{x}_0}^\vartheta}(1)$ and $\int g^2 dN_s^\gamma = O_{\mathbb{P}_{\mathbf{x}_0}^\vartheta}(1)$. We have

$$\begin{aligned} & \mathbb{E}_{\mathbf{x}_0}^\vartheta \int (f - g)^2 dN_s^\gamma \\ &= \mathbb{E}_{\mathbf{x}_0}^\vartheta \int_0^{tN} \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 - \frac{h^T \nabla_1 \gamma(\vartheta, \mathbf{X}_{s^-})}{\sqrt{N} \gamma(\vartheta, \mathbf{X}_{s^-})} \right)^2 dN_s^\gamma \\ &= \mathbb{E}_{\mathbf{x}_0}^\vartheta \int_0^{tN} \left(\frac{\gamma(\vartheta_N, \mathbf{X}_s)}{\gamma(\vartheta, \mathbf{X}_s)} - 1 - \frac{h^T \nabla_1 \gamma(\vartheta, \mathbf{X}_s)}{\sqrt{N} \gamma(\vartheta, \mathbf{X}_s)} \right)^2 \mathbf{1}_{\gamma(\vartheta, \mathbf{X}_s) > 0} \gamma(\vartheta, \mathbf{X}_s) ds, \end{aligned} \quad (\text{D.15})$$

which is exactly $\mathbb{E}_{\mathbf{x}_0}^\vartheta(\langle S^N \rangle_t)$, see (D.9). We have already proven that this term tends to 0, so $\int (f - g)^2 dN_s^\gamma = o_{\mathbb{P}_{\mathbf{x}_0}^\vartheta}(1)$. Now, for any $\eta > 0$, by the Markov's inequality, we can choose $M > 0$ such that

$$\mathbb{P}_{\mathbf{x}_0}^\vartheta \left(\int g^2 dN_s^\gamma > M \right) \leq \frac{1}{M} \mathbb{E}_{\mathbf{x}_0}^\vartheta \left(\frac{1}{N} \int_0^{tN} \frac{(h^T \nabla_1 \gamma(\vartheta, \mathbf{X}_s))^2}{\gamma(\vartheta, \mathbf{X}_s)} ds \right) \leq \frac{t}{M} C(\vartheta, h) < \eta,$$

where $C(\vartheta, h)$ is a positive upper-bound deduced from condition $\mathbf{A}(\gamma)$. This proves that $\int g^2 dN_s^\gamma = O_{\mathbb{P}_{\mathbf{x}_0}^\vartheta}(1)$ and completes the proof that R_2 tends to 0 in probability.

It remains to address the convergence of R_1 . For any $\epsilon > 0$, we have $R_1 = R_{11} + R_{12}$ with

$$R_{11} = \int_0^{tN} \varphi \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 \right) \mathbf{1}_{\left\{ \left| \frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 \right| > \epsilon \right\}} dN_s^\gamma, \quad (\text{D.16})$$

$$R_{12} = \int_0^{tN} \varphi \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 \right) \mathbf{1}_{\left\{ \left| \frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 \right| \leq \epsilon \right\}} dN_s^\gamma. \quad (\text{D.17})$$

For R_{11} , note that for any $a > 0$,

$$\mathbb{P}_{\mathbf{x}_0}^\vartheta (R_{11} > a) \leq \mathbb{P}_{\mathbf{x}_0}^\vartheta \left(\int_0^{tN} \mathbf{1}_{\left\{ \left| \frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 \right| > \epsilon \right\}} dN_s^\gamma > b \right), \quad (\text{D.18})$$

for any $b < 1$. This is because when $b < 1$,

$$\begin{aligned} & \left(\int_0^{tN} \mathbf{1}_{\left\{ \left| \frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 \right| > \epsilon \right\}} dN_s^\gamma \leq b \right) \\ & \implies \left(\text{for all } \gamma\text{-jump time } T_i \leq tN, \mathbf{1}_{\left\{ \left| \frac{\gamma(\vartheta_N, \mathbf{X}_{T_i^-})}{\gamma(\vartheta, \mathbf{X}_{T_i^-})} - 1 \right| > \epsilon \right\}} = 0 \right) \end{aligned}$$

which implies that $R_{11} = 0 \leq a$. Moreover

$$\mathbb{1}\left\{\left|\frac{\gamma(\vartheta_N, \mathbf{x})}{\gamma(\vartheta, \mathbf{x})} - 1\right| > \epsilon\right\} \leq \mathbb{1}\left\{\left|\frac{\gamma(\vartheta_N, \mathbf{x})}{\gamma(\vartheta, \mathbf{x})} - 1 - \frac{h^T \nabla_1 \gamma(\vartheta, \mathbf{x})}{\sqrt{N} \gamma(\vartheta, \mathbf{x})}\right| > \frac{\epsilon}{2}\right\} + \mathbb{1}\left\{\left|\frac{h^T \nabla_1 \gamma(\vartheta, \mathbf{x})}{\sqrt{N} \gamma(\vartheta, \mathbf{x})}\right| > \frac{\epsilon}{2}\right\}.$$

Using this and the fact that for any z , $\mathbb{1}\{|z| > \frac{\epsilon}{2}\} \leq 4\epsilon^{-2} z^2 \mathbb{1}\{|z| > \frac{\epsilon}{2}\}$, we get

$$\begin{aligned} & \int_0^{tN} \mathbb{1}\left\{\left|\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1\right| > \epsilon\right\} dN_s^\gamma \\ & \leq 4\epsilon^{-2} \int_0^{tN} \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 - \frac{h^T \nabla_1 \gamma(\vartheta, \mathbf{X}_{s^-})}{\sqrt{N} \gamma(\vartheta, \mathbf{X}_{s^-})}\right)^2 dN_s^\gamma \\ & \quad + 4\epsilon^{-2} \int_0^{tN} \left(\frac{h^T \nabla_1 \gamma(\vartheta, \mathbf{X}_{s^-})}{\sqrt{N} \gamma(\vartheta, \mathbf{X}_{s^-})}\right)^2 \mathbb{1}\left\{\left|\frac{h^T \nabla_1 \gamma(\vartheta, \mathbf{X}_{s^-})}{\sqrt{N} \gamma(\vartheta, \mathbf{X}_{s^-})}\right| > \frac{\epsilon}{2}\right\} dN_s^\gamma. \end{aligned}$$

The first term is exactly $4\epsilon^{-2} \int (f - g)^2 dN_s^\gamma$, as already studied in (D.15), and the second term is $4\epsilon^{-2} A_2$ (up to $\epsilon/2$ instead of ϵ), see (D.13). Both terms tend to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability whatever the value of $\epsilon > 0$. This proves that the right-hand side term in (D.18) tends to 0 for any $b < 1$, which yields the convergence of R_{11} in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability to 0.

For R_{12} given by (D.17), we choose $\epsilon < \frac{1}{2}$ and since $|\varphi(x)| \leq 2|x|^3$ for $|x| \leq \frac{1}{2}$, we have

$$\begin{aligned} R_{12} & \leq \int_0^{tN} 2 \left| \frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1 \right|^3 \mathbb{1}\left\{\left|\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1\right| \leq \epsilon\right\} dN_s^\gamma \\ & \leq 2\epsilon \int_0^{tN} \left(\frac{\gamma(\vartheta_N, \mathbf{X}_{s^-})}{\gamma(\vartheta, \mathbf{X}_{s^-})} - 1\right)^2 dN_s^\gamma. \end{aligned}$$

This last term is $2\epsilon(R_2 + \int g^2 dN_s^\gamma)$ if we use as previously the notation $R_2 = \int (f^2 - g^2) dN_s^\gamma$ for R_2 given by (D.14). We already know that $U := 2(R_2 + \int g^2 dN_s^\gamma)$ is a $O_{\mathbb{P}_{\mathbf{x}_0}^\vartheta}(1)$ as $N \rightarrow \infty$. This means that for any $\eta > 0$, there exists M such that for N sufficiently large $\mathbb{P}_{\mathbf{x}_0}^\vartheta(U > M) < \eta$. This implies that for any $a > 0$ and $\eta > 0$, we can choose ϵ small enough ($\epsilon < a/M$) so that $\mathbb{P}_{\mathbf{x}_0}^\vartheta(\epsilon U > a) < \mathbb{P}_{\mathbf{x}_0}^\vartheta(U > M) < \eta$. So for any $a > 0$ and $\eta > 0$, we can choose ϵ so that $\mathbb{P}_{\mathbf{x}_0}^\vartheta(R_{12} > a) < \eta$ for N large enough. The same inequality holds true for R_{11} since we have proven that R_{11} tends to 0 in probability for any $\epsilon > 0$. So $R_1 = R_{11} + R_{12}$ tends to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability, which concludes the proof.

D.2 Proof of Lemma 2

The proof follows the same scheme as the proof of Lemma 1. Consider

$$Rem_{N,\vartheta,h}^{k_\gamma}(t) = \Lambda_{k_\gamma}^{\vartheta_N/\vartheta}(tN) - h^T M_{k_\gamma}^{N,\vartheta}(t) + \frac{1}{2} h^T \langle M_{k_\gamma}^{N,\vartheta} \rangle(t) h,$$

where $M_{k_\gamma}^{N,\vartheta}(t)$ is given by (3.10). The fact that $M_{k_\gamma}^{N,\vartheta}$ is a \mathcal{F}_{tN} -martingale is a consequence of Lemma 4. We deduce that

$$\begin{aligned} [M_{k_\gamma}^{N,\vartheta}](t) &= \frac{1}{N} \int_0^{tN} \frac{\nabla_1 k_\gamma (\nabla_1 k_\gamma)^T}{k_\gamma^2}(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s) dN_s^\gamma, \\ \langle M_{k_\gamma}^{N,\vartheta} \rangle(t) &= \frac{1}{N} \int_0^{tN} \int_E \mathbf{1}_{k_\gamma > 0} \frac{\nabla_1 k_\gamma (\nabla_1 k_\gamma)^T}{k_\gamma}(\vartheta, \mathbf{X}_s, \mathbf{y}) \nu_\gamma(\mathbf{X}_s, d\mathbf{y}) \gamma(\vartheta, \mathbf{X}_s) ds. \end{aligned}$$

Starting from (D.3) and using the above representations we can decompose $Rem_{N,\vartheta,h}^{k_\gamma}(t)$ as in the proof of Lemma 1, that is $Rem_{N,\vartheta,h}^{k_\gamma}(t) = S_t^N + A_t^N + R_t^N$, where

$$\begin{aligned} S_t^N &= \int_0^{tN} \left(\frac{k_\gamma(\vartheta_N, \mathbf{X}_{s^-}, \mathbf{X}_s)}{k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)} - 1 - \frac{h^T \nabla_1 k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)}{\sqrt{N} k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)} \right) dN_s^\gamma \\ &\quad + \frac{1}{\sqrt{N}} \int_0^{tN} \int_E h^T \nabla_1 k_\gamma(\vartheta, \mathbf{X}_s, \mathbf{y}) \nu_\gamma(\mathbf{X}_s, d\mathbf{y}) \gamma(\vartheta, \mathbf{X}_s) ds, \end{aligned}$$

$$\begin{aligned} A_t^N &= \frac{1}{2N} \int_0^{tN} \int_E \mathbf{1}_{k_\gamma > 0} \frac{(h^T \nabla_1 k_\gamma)^2}{k_\gamma}(\vartheta, \mathbf{X}_s, \mathbf{y}) \nu_\gamma(\mathbf{X}_s, d\mathbf{y}) \gamma(\vartheta, \mathbf{X}_s) ds \\ &\quad - \frac{1}{2N} \int_0^{tN} \frac{(h^T \nabla_1 k_\gamma)^2}{k_\gamma^2}(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s) dN_s^\gamma, \end{aligned}$$

$$\begin{aligned} R_t^N &= \int_0^{tN} \left(\log \left(\frac{k_\gamma(\vartheta_N, \mathbf{X}_{s^-}, \mathbf{X}_s)}{k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)} \right) - \frac{k_\gamma(\vartheta_N, \mathbf{X}_{s^-}, \mathbf{X}_s)}{k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)} + 1 \right. \\ &\quad \left. + \frac{1}{2N} \frac{(h^T \nabla_1 k_\gamma)^2}{k_\gamma^2}(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s) \right) dN_s^\gamma. \end{aligned}$$

We show that each of these three terms tends to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability as in the proof of Lemma 1.

First step : We have

$$[S_t^N] = \int_0^{tN} \left(\frac{k_\gamma(\vartheta_N, \mathbf{X}_{s^-}, \mathbf{X}_s)}{k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)} - 1 - \frac{h^T \nabla_1 k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)}{\sqrt{N} k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)} \right)^2 dN_s^\gamma$$

so that by Lemma 4 in Section D.4,

$$\begin{aligned} \langle S^N \rangle_t &= \int_0^{tN} \int_E \left(\frac{k_\gamma(\vartheta_N, \mathbf{X}_s, \mathbf{y})}{k_\gamma(\vartheta, \mathbf{X}_s, \mathbf{y})} - 1 - \frac{h^T \nabla_1 k_\gamma(\vartheta, \mathbf{X}_s, \mathbf{y})}{\sqrt{N} k_\gamma(\vartheta, \mathbf{X}_s, \mathbf{y})} \right)^2 \\ &\quad \times k_\gamma(\vartheta, \mathbf{X}_s, \mathbf{y}) \nu_\gamma(\mathbf{X}_s, d\mathbf{y}) \gamma(\vartheta, \mathbf{X}_s) ds. \end{aligned}$$

Using the same inequalities as in (D.10) we obtain that

$$\langle S^N \rangle_t \leq \frac{\|h\|^2}{N} \int_0^{tN} f_\vartheta^{k_\gamma} \left(\mathbf{X}_s, \frac{\|h\|}{\sqrt{N}} \right) \gamma(\vartheta, \mathbf{X}_s) ds$$

where $f_\vartheta^{k_\gamma}$ is given by (3.7). We can then deduce that $\mathbb{E}_{\mathbf{x}_0}^\vartheta(\langle S^N \rangle_t)$ tends to 0 by use of Conditions **B**(k_γ, γ), **(H)** and the dominated convergence theorem, as in the proof of Lemma 1.

Second step : The proof is identical to the second step of the proof of Lemma 1. It consists in splitting the integrals in A_t^N according to whether $\left| \frac{h^T \nabla_1 k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)}{k_\gamma(\vartheta, \mathbf{X}_{s^-}, \mathbf{X}_s)} \right| > \epsilon \sqrt{N}$ or not. The convergence of A_t^N is then proved similarly, by use of Conditions **A**(k_γ, γ), **(H)** and the dominated convergence theorem.

Third step : The proof follows exactly the same lines as the proof of the third step of Lemma 1 where $\gamma(\cdot, \cdot)$ is replaced by $k_\gamma(\cdot, \cdot, \cdot)$ and Condition **A**(k_γ, γ) is used instead of Condition **A**(γ).

D.3 Proof of Lemma 3

Remember that $\mathbf{X}_t = \{(\mathbf{Z}_{t-T_i}^{(i)}, \mathbf{m}^{(i)})\}$ for all $t \in [T_i, T_{i+1}[$, and that $L(\vartheta, \mathbf{X}_{[T_i, T_{i+1}[})$ is a shortcut for $L(\mathbf{Z}_{[0, T_{i+1}-T_i]}^{(i)}, \mathbf{m}^{(i)})$ given by (3.5). Accordingly, we have

$$\begin{aligned} \log L(\vartheta, \mathbf{X}_{[T_i, T_{i+1}[}) &= \int_{T_i}^{T_{i+1}} \bar{v}_n(\vartheta, \mathbf{X}_s)^T a_n^{-1}(\mathbf{X}_s) d\mathbf{Z}_{s-T_i}^{(i)} \\ &\quad - \frac{1}{2} \int_{T_i}^{T_{i+1}} \bar{v}_n(\vartheta, \mathbf{X}_s)^T a_n^{-1}(\mathbf{X}_s) \bar{v}_n(\vartheta, \mathbf{X}_s) ds, \end{aligned}$$

where $\bar{v}_n(\vartheta, \mathbf{X}_s)$ stands for $\bar{v}_n(\vartheta, \mathbf{Z}_{s-T_i}^{(i)}, \mathbf{m}^{(i)})$, with $n = n(\mathbf{X}_s)$, and similarly for $a_n^{-1}(\mathbf{X}_s)$. Then we have

$$\begin{aligned} \log \frac{L(\vartheta_N, \mathbf{X}_{[T_i, T_{i+1}]})}{L(\vartheta, \mathbf{X}_{[T_i, T_{i+1}]})} &= \int_{T_i}^{T_{i+1}} (\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s))^T a_n^{-1}(\mathbf{X}_s) (d\mathbf{Z}_{s-T_i}^{(i)} - \bar{v}_n(\vartheta, \mathbf{X}_s) ds) \\ &\quad - \frac{1}{2} \int_{T_i}^{T_{i+1}} (\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s))^T a_n^{-1}(\mathbf{X}_s) (\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s)) ds. \end{aligned}$$

Consequently, we may rewrite (D.4) as

$$\begin{aligned} \Lambda_L^{\vartheta_N/\vartheta}(t) &= \sum_{i \geq 0} \int_0^t \mathbb{1}_{[T_i, T_{i+1}[}(s) (\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s))^T a_n^{-1}(\mathbf{X}_s) dM_{s-T_i}^{(i)} \\ &\quad - \frac{1}{2} \int_0^t (\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s))^T a_n^{-1}(\mathbf{X}_s) (\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s)) ds. \end{aligned} \quad (\text{D.19})$$

where $M_s^{(i)} = 0$ if $s < 0$ and

$$M_s^{(i)} = \mathbf{Z}_s^{(i)} - \mathbf{Z}_0^{(i)} - \int_0^s \bar{v}_n(\vartheta, \mathbf{Z}_u^{(i)}) du, \quad s \geq 0. \quad (\text{D.20})$$

We know that $\mathbf{Z}^{(i)}$ is the solution of $M^{(i)}$ given by (3.4), where $(\mathbf{z}^{(i)}, \mathbf{m}^{(i)})$ are such that $\mathbf{X}_{T_i} = \{(\mathbf{z}^{(i)}, \mathbf{m}^{(i)})\}$ and $n = n(\mathbf{X}_{T_i})$. This implies that $(M_s^{(i)})$ is a martingale with respect to $\mathcal{F}_t^{\mathbf{Z}^{(i)}}$, the natural filtration associated to $\mathbf{Z}^{(i)}$. Note that $(M_s^{(i)})$ is independent from T_i and that \mathcal{F}_t is generated by :

$$\left\{ T_{N_t}, \mathbf{X}_{T_{N_t}}, \mathbf{Z}_{[0, t-T_{N_t}]}^{(N_t)}, (T_i, \mathbf{X}_{T_i}, \mathbf{Z}_{[0, T_{i+1}-T_i]}^{(i)})_{i=1, \dots, N_t-1} \right\}.$$

This implies that for $T_i \leq t < T_{i+1}$, if $s \in [T_i, t]$,

$$\mathbb{E}(M_{t-T_i}^{(i)} | \mathcal{F}_s) = \mathbb{E}(M_{t-T_i}^{(i)} | T_i, \mathcal{F}_{s-T_i}^{\mathbf{Z}^{(i)}}) = M_{s-T_i}^{(i)},$$

while if $s < T_i$, $\mathbb{E}(M_{t-T_i}^{(i)} | \mathcal{F}_s) = 0$.

Using these properties, we can easily verified that $M_L^{N, \vartheta}$ given by (3.11) is a \mathcal{F}_{tN} -martingale. Similarly, since $\langle M^{(i)} \rangle_s = \int_0^s a_n(\mathbf{Z}_u^{(i)}) du$, we have

$$\begin{aligned} \langle M_L^{N, \vartheta} \rangle(t) &= \frac{1}{N} \sum_{i \geq 0} \int_0^{tN} \mathbb{1}_{[T_i, T_{i+1}[}(s) (\nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s))^T a_n^{-1}(\mathbf{X}_s) a_n(\mathbf{X}_s) a_n^{-1}(\mathbf{X}_s) \nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s) ds \\ &= \frac{1}{N} \int_0^{tN} (\nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s))^T a_n^{-1}(\mathbf{X}_s) \nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s) ds. \end{aligned}$$

From this representation, (3.11) and (D.19), we have

$$Rem_{N,\vartheta,h}^L(t) = \Lambda_L^{\vartheta_N/\vartheta}(tN) - h^T M_L^{N,\vartheta}(t) + \frac{1}{2} h^T \langle M_L^{N,\vartheta} \rangle(t) h = S_t^N - \frac{1}{2} R_t^N,$$

where

$$S_t^N = \sum_{i \geq 0} \int_0^{tN} \mathbb{1}_{[T_i; T_{i+1}[}(s) \left(\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s) - \nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s) \frac{h}{\sqrt{N}} \right)^T a_n^{-1}(\mathbf{X}_s) dM_{s-T_i}^{(i)},$$

and

$$R_t^N = \int_0^{tN} \left((\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s))^T a_n^{-1}(\mathbf{X}_s) (\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s)) - \frac{1}{N} h^T (\nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s))^T a_n^{-1}(\mathbf{X}_s) \nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s) h \right) ds.$$

Let us prove that both S_t^N and R_t^N tend to 0 in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability. For S_t^N , using the Markov inequality and the fact that $\mathbb{E}_{\mathbf{x}_0}^\vartheta((S_t^N)^2) = \mathbb{E}_{\mathbf{x}_0}^\vartheta(\langle S_t^N \rangle)$, this boils down to proving that $\mathbb{E}_{\mathbf{x}_0}^\vartheta(\langle S_t^N \rangle)$ tends to 0. We have, since $a_n = \bar{\sigma}_n \bar{\sigma}_n^T$,

$$\begin{aligned} \langle S_t^N \rangle &= \int_0^{tN} \left(\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s) - \nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s) \frac{h}{\sqrt{N}} \right)^T a_n^{-1}(\mathbf{X}_s) \\ &\quad \times \left(\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s) - \nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s) \frac{h}{\sqrt{N}} \right) ds \\ &= \int_0^{tN} \left\| \bar{\sigma}_n(\mathbf{X}_s)^{-1} \left(\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s) - \nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s) \frac{h}{\sqrt{N}} \right) \right\|^2 ds. \end{aligned}$$

Using the same argument as for (D.10), we get

$$\langle S_t^N \rangle \leq \frac{\|h\|^2}{N} \int_0^{tN} \left\| \bar{\sigma}_n(\mathbf{X}_s)^{-1} \right\|^2 f_\vartheta^{\text{move}} \left(\mathbf{X}_s, \frac{\|h\|}{\sqrt{N}} \right) ds.$$

where $f_\vartheta^{\text{move}}$ is given by (3.8). We then deduce that $\mathbb{E}_{\mathbf{x}_0}^\vartheta(\langle S_t^N \rangle)$ tends to 0 thanks to the dominated convergence theorem, using Conditions \mathbf{B}_{move} and (\mathbf{H}) as in the first step of the proof of Lemma 1.

Concerning R_t^N , note that

$$R_t^N = \int_0^{tN} \left\| \bar{\sigma}_n(\mathbf{X}_s)^{-1} (\bar{v}_n(\vartheta_N, \mathbf{X}_s) - \bar{v}_n(\vartheta, \mathbf{X}_s)) \right\|^2 - \frac{1}{N} \left\| \bar{\sigma}_n(\mathbf{X}_s)^{-1} \nabla_1 \bar{v}_n(\vartheta, \mathbf{X}_s) h \right\|^2 ds,$$

that reads $\int (\|f\|^2 - \|g\|^2) ds$ with obvious notation, so that using the same argument as in the third step of the proof of Lemma 1 (term R_2), it suffices to show that $\int \|f - g\|^2 ds = o(1)$ and $\int \|g\|^2 ds = O(1)$ in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability. But the former is exactly $\langle S_t^N \rangle$ studied above, that has been proven to converge to 0 in probability. And the expectation of the latter is bounded by use of Condition \mathbf{A}_{move} , proving that it is $O(1)$ in $\mathbb{P}_{\mathbf{x}_0}^\vartheta$ -probability. This concludes the proof.

D.4 A useful martingale

The following result is useful in the proof of Theorem 2 in Section D. It is proven in Lemma 54 of [Man23], the BDMM being a particular case of a jump-move process (see the proof of Theorem 1 in Section C). It is also stated under a slightly different setting in Proposition 3.3 (b) of [Löc02b].

Lemma 4. *Let $(\mathbf{X}_t)_{t \geq 0}$ be a BDMM process on E , let g be a measurable bounded function defined on $E \times E$ and introduce for any $t > 0$*

$$M_t^* = \int_0^t g(\mathbf{X}_{s-}, \mathbf{X}_s) dN_s^\gamma - \int_0^t \gamma(\mathbf{X}_s) \int_E g(\mathbf{X}_s, \mathbf{y}) k_\gamma(\mathbf{X}_s, \mathbf{y}) \nu_\gamma(\mathbf{X}_s, d\mathbf{y}) ds.$$

If $N_t^\gamma < \infty$ for any $t \geq 0$, then $(M_t^)_{t \geq 0}$ is a \mathcal{F}_t -martingale.*

ERGODICITY OF THE BDMM PROCESS

The following proposition provides some conditions ensuring the hypothesis **(H)**, namely non-explosion and geometric ergodicity of the BDMM process. It is based on the study conducted in [LLGM22] for birth-death-move processes (without mutations). The arguments in presence of mutations are similar and sketched below. The conditions include the technical assumption that the process is Feller. This property is discussed in [LLGM22] and proved for several examples of transition kernels that include all examples of this paper, provided the underlying spatial state Λ is compact. The other conditions deal with the intensity functions of the jumps. First, the total intensity function α is assumed to be bounded, to avoid explosion of the process. Second, the death intensity function δ must compensate in a proper way the birth intensity function β . Notably, the ergodic properties do not depend on the inter-jump move process, neither on the mutation dynamics. We introduce the following notation :

$$\beta_n = \sup_{\mathbf{x} \in E_n} \beta(\mathbf{x}), \quad \delta_n = \inf_{\mathbf{x} \in E_n} \delta(\mathbf{x}) \quad \text{and} \quad \alpha_n = \beta_n + \delta_n. \quad (\text{E.1})$$

In fact the conditions of geometric ergodicity below are exactly the same as the conditions of ergodicity for a simple birth-death process $(\eta_t)_{t \geq 0}$ on \mathbb{N} with birth rate (β_n) and death rate (δ_n) , as established in [KM57]. This is because to address the ergodic properties of a BDMM $(\mathbf{X}_t)_{t \geq 0}$, we construct a coupling between $(\mathbf{X}_t)_{t \geq 0}$ and $(\eta_t)_{t \geq 0}$, in such a way that $\eta_t = 0$ implies $\mathbf{X}_t = \emptyset$. Consequently \emptyset becomes a positive recurrent state for $(\mathbf{X}_t)_{t \geq 0}$ whenever 0 is a positive recurrent state for $(\eta_t)_{t \geq 0}$, which implies the geometric ergodicity of $(\mathbf{X}_t)_{t \geq 0}$.

Let $(\mathbf{X}_t)_{t \geq 0}$ be a Feller BDMM process with a bounded intensity function α .

Suppose that $\delta_n > 0$ for all $n \geq 1$ and one of the following condition holds :

$$(i) \text{ there exists } n_0 \geq 1 \text{ such that } \beta_n = 0 \text{ for any } n \geq n_0, \quad (\text{E.2})$$

$$(ii) \beta_n > 0 \text{ for all } n \geq 1, \sum_{n=2}^{\infty} \frac{\beta_1 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n} < \infty \text{ and } \sum_{n=1}^{\infty} \frac{\delta_1 \cdots \delta_n}{\beta_1 \cdots \beta_n} = \infty, \quad (\text{E.3})$$

where β_n and δ_n are defined by (E.1). Assume moreover that

$$\sum_{n=2}^{\infty} \sqrt{\frac{\beta_1 \cdots \beta_{n-1}}{\delta_1 \cdots \delta_n}} < \infty \quad \text{and} \quad \exists N \geq 0, \text{ s.t. } \forall n \geq N, \beta_n \leq \delta_{n+1}.$$

Then **(H)** is satisfied.

As mentioned above, following [Pre75] and [LLGM22], the proof of this result is based on a coupling between $(\mathbf{X}_t)_{t \geq 0}$ and $(\eta_t)_{t \geq 0}$, where the latter is a simple birth-death process with birth and death rates given by (E.1). We detail below how we construct the coupled process $\check{C}_t = (\mathbf{X}_t, \eta_t)$. This is a straightforward generalisation of the construction in [LLGM22], where we account for the presence of mutations. Specifically, \check{C} is a jump move process on $E \times \mathbb{N}$ with intensity function $\check{\alpha}$ and transition kernel \check{K} defined as follows. The intensity function $\check{\alpha} : E \times \mathbb{N} \rightarrow \mathbb{R}_+$ is given by

$$\check{\alpha}(\mathbf{x}, n) = \begin{cases} \alpha(\mathbf{x}) + \alpha_n & \text{if } \mathbf{x} \in E_m, m \neq n, \\ \beta_n + \delta(\mathbf{x}) + \tau(\mathbf{x}) & \text{if } \mathbf{x} \in E_n. \end{cases}$$

Letting K be the kernel given by (C.1), the transition kernel \check{K} takes the form, for any $A \subset E$:

1. If $\mathbf{x} \in E_m, m \neq n$:

$$\check{K}((\mathbf{x}, n); A \times \{n\}) = \frac{\alpha(\mathbf{x})}{\check{\alpha}(\mathbf{x}, n)} K(\mathbf{x}, A);$$

$$\check{K}((\mathbf{x}, n); \{\mathbf{x}\} \times \{n+1\}) = \frac{\beta_n}{\check{\alpha}(\mathbf{x}, n)};$$

$$\check{K}((\mathbf{x}, n); \{\mathbf{x}\} \times \{n-1\}) = \frac{\delta_n}{\check{\alpha}(\mathbf{x}, n)}.$$

2. If $\mathbf{x} \in E_n$:

$$\check{K}((\mathbf{x}, n); A \times \{n+1\}) = \frac{\beta(\mathbf{x})}{\check{\alpha}(\mathbf{x}, n)} K_{\beta}(\mathbf{x}, A);$$

$$\begin{aligned}\check{K}((\mathbf{x}, n); \{\mathbf{x}\} \times \{n+1\}) &= \frac{\beta_n - \beta(\mathbf{x})}{\check{\alpha}(\mathbf{x}, n)}; \\ \check{K}((\mathbf{x}, n); A \times \{n-1\}) &= \frac{\delta_n}{\check{\alpha}(\mathbf{x}, n)} K_\delta(\mathbf{x}, A); \\ \check{K}((\mathbf{x}, n); A \times \{n\}) &= \frac{\delta(\mathbf{x}) - \delta_n}{\check{\alpha}(\mathbf{x}, n)} K_\delta(\mathbf{x}, A) + \frac{\tau(\mathbf{x})}{\check{\alpha}(\mathbf{x}, n)} K_\tau(\mathbf{x}, A).\end{aligned}$$

The inter-jump move process of \check{C} is in turn a simple independent coupling between the move \mathbf{Y} of \mathbf{X} and a constant move on \mathbb{N} (i.e. $y_t = y_0$ for all $t \geq 0$).

The fact that the above construction is a proper coupling, in the sense that \mathbf{X} and η do constitute the marginal distributions of \check{C} , can be proven exactly as in [LLGM22] for the case without mutations. On the other hand, the key point is that if at some point, $\check{C}_t = (\mathbf{X}_t, \eta_t)$ is such that $n(\mathbf{X}_t) \leq \eta_t$, then by the above construction, $n(\mathbf{X}_s) \leq \eta_s$ for all $s \geq t$, almost surely. By this property, $\eta_t = 0$ implies $\mathbf{X}_t = \emptyset$ and the geometric ergodic conditions for η stated in the proposition, coming from [KM57], are sufficient for the geometric ergodicity of \mathbf{X} . The rigorous proof of these claims can be found in [LLGM22].

PARTIE III

**Analyse du mouvement des
trajectoires et détection des
instants de changement de régime.**

MODÈLES DE MARKOV À ÉTATS CACHÉS ET ALGORITHME EM POUR CLASSER DES TRAJECTOIRES

Afin de complètement caractériser les estimateurs paramétriques des différents éléments du processus BDMM, il faut être capable de déterminer les instants de saut qui sont des changements de marque. En effet, les instants de saut associés à des naissances ou des morts sont faciles à identifier lorsque l'on observe des particules dans un domaine donné : il suffit de porter son attention sur les points de temps où le nombre de particules change. En revanche, les instants de changement de marque sont plus difficiles à estimer. Nous allons nous placer dans le cas où la marque associée à une particule désigne son type de mouvement. Notre but est alors de déterminer quel mouvement est suivi par chacune des particules et à quels moments certaines particules changent de type de mouvement. Dans ce chapitre, nous nous plaçons dans une situation où le mouvement observé sur une trajectoire donnée dépend uniquement de la marque de la particule, indépendamment des autres trajectoires. De plus, nous considérons aussi que lorsqu'une particule change de marque, ce changement se fait indépendamment des positions et des marques des autres particules. Nous allons, sous ces hypothèses, donner une méthode permettant d'estimer le mouvement suivi par chaque particule ainsi que les instants de changement de mouvement.

Dans la littérature, beaucoup de méthodes cherchent à classer les mouvements de trajectoires supposées toutes indépendantes. Ce chapitre commence par un rapide état de l'art sur le sujet, puis se poursuit par la présentation de notre méthode de classification, basée sur des modèles de Markov à états cachés et un algorithme EM pour l'estimation des paramètres. Enfin, nous traitons le cas particulier où l'ensemble de modèles de mouvement possibles est $\mathbb{M} = \{\text{brownien, processus d'Ornstein-Uhlenbeck, brownien dirigé}\}$.

5.1 État de l’art

Actuellement, beaucoup de méthodes utilisent le déplacement quadratique moyen (MSD : *Mean Square Displacement*) [GLGW13, QSE91] pour classer les mouvements de diffusion. Le MSD d’une particule au temps t est défini par $MSD(t) = n^{-1} \sum_{i=0}^{n-1} \|z_{t_{k+1}} - z_{t_k}\|^2$ où z_{t_0}, \dots, z_{t_n} sont les positions successives de la particule considérée aux temps $0 = t_0 < \dots < t_n = t$. Le calcul du MSD d’une particule permet de classer son mouvement en trois catégories : mouvement brownien lorsque le MSD est proportionnel à t , mouvement sous-diffusif lorsque le MSD est proportionnel à t^α avec $0 < \alpha < 1$ et super-diffusif lorsque $\alpha > 1$. Le but de ces méthodes est d’estimer le coefficient α afin d’en déduire le type de mouvement sous-jacent [BKV18]. La précision du résultat est fortement dépendante du nombre de points acquis au cours de l’observation, ce qui engendre une mauvaise analyse des trajectoires courtes. Aussi, l’étude du MSD se faisant sur les trajectoires complètes, la plupart de ces méthodes ne permettent pas de détecter le passage d’un type de mouvement à un autre au cours du déplacement d’une particule.

Afin de détecter les moments de changement de type de mouvement au sein d’une même trajectoire, des nouvelles méthodes ont été mises en place. Tout d’abord, [BVK19] présente une extension de la procédure de test non-paramétrique décrite dans [BKV18], pour détecter les instants où une particule change de mode de diffusion (diffusion brownienne, sous-diffusion, super-diffusion). Une statistique de test est calculée sur des fenêtres glissantes le long de la trajectoire. Malheureusement, les résultats de cette méthode sont fortement dépendants du choix de la taille de la fenêtre glissante.

Par ailleurs, la plupart des méthodes existantes se basent sur des représentations des états par une chaîne de Markov à états cachés pour détecter des changements de mouvement sur une trajectoire. Ces méthodes n’envisagent que des mouvements browniens et des mouvement browniens dirigés. Dans [CAV+10], les auteurs se servent du MSD, ou plus précisément du CSD (*Cumulative Square Displacement*) pour trouver les paramètres optimaux de deux types de mouvement possibles (un mouvement brownien et un mouvement diffusif) avant de retrouver le type de mouvement sur chaque morceau de trajectoire grâce aux modèles de Markov à états cachés. Dans [PLUE13], les auteurs classent des trajectoires selon plusieurs mouvements

browniens associés à des coefficients de diffusion différents. Il déterminent d'abord avec une approche bayésienne variationnelle le nombre de mouvements browniens adéquats pour décrire les données observées, puis les paramètres optimaux, avant d'utiliser les modèles de Markov à états cachés pour classer les trajectoires. Enfin, [MBP⁺15] présente une approche plus élaborée, pour laquelle les trajectoires observées sont supposées provenir d'un mélange de K mouvements browniens et K_V mouvements browniens dirigés. Les auteurs utilisent une méthode bayésienne pour déterminer ces nombres K et K_V optimaux pour les trajectoires observées. Ensuite, ils utilisent une méthode d'échantillonnage MCMC (Markov chain Monte Carlo) pour déterminer la valeur du maximum de vraisemblance, qui leur permettra, grâce aux modèles de Markov à états cachés, d'estimer le type de mouvement pour chaque point de la trajectoire.

D'autres méthodes basées sur l'apprentissage profond (Deep Learning) [ASP⁺19, MGVM⁺21] produisent des résultats bien meilleurs : l'idée est d'entraîner un réseau neuronal avec des trajectoires simulées comportant des sauts de mouvement. Une fois le réseau entraîné, on l'applique sur des données afin de découper chaque trajectoire en segments qui présentent le même type de mouvement. Ensuite, chaque segment est analysé par différentes méthodes pour en déduire ses paramètres caractéristiques. Le gros inconvénient de ces méthodes d'apprentissage automatique est qu'elles fonctionnent comme une boîte noire et on peut donc préférer des estimateurs basés sur des statistiques explicites afin d'en tirer profit pour en déduire d'autres informations sur la diffusion étudiée.

La méthode que nous avons développée dans ce manuscrit utilise les modèles de Markov à états cachés pour estimer le type de mouvement pour chaque point de chaque trajectoire, une fois les paramètres qui caractérisent ces mouvements déterminés. Nous utilisons un algorithme EM (Expectation-Maximization) pour déterminer les paramètres du modèle. Notre approche permet de gérer un grand nombre de trajectoires. L'annexe H présente une comparaison entre les résultats obtenus par notre méthode et ceux obtenus par la méthode présentée dans [MBP⁺15], qui classe les segments d'une trajectoire suivant deux types de mouvement (brownien dirigé ou brownien). Cette dernière inclut la possibilité qu'une même trajectoire puisse suivre plusieurs types de mouvements browniens et plusieurs types de mouvements dirigés, avec des coefficients de diffusion et des drifts différents. L'intérêt de cette méthode

est de ne pas supposer connu le nombre de types de mouvement et de renvoyer le nombre le plus probable en plus des caractéristiques des mouvements correspondants. Cette méthode présente l'inconvénient majeur de ne plus fonctionner dès qu'il y a plus de deux trajectoires à analyser. De plus, pour l'analyse d'une seule trajectoire, le temps de calcul nécessaire pour analyser une trajectoire est relativement long, ce qui rendrait difficile son adaptation pour étudier d'un grand nombre de trajectoires.

5.2 Résolution du problème

Dans ce chapitre, on suppose avoir relevé les positions du processus BDMM à intervalles de temps discrets, notés t_0, \dots, t_n , chacun espacé de $\Delta = |t_{k+1} - t_k|$. On suppose que l'on sait suivre chaque trajectoire, c'est-à-dire que le jeu de données a pu être traité par un outil de tracking qui a permis de suivre chaque particule et ainsi de connaître les trajectoires du système. On connaît donc le nombre s de trajectoires présentes entre les deux temps d'étude t_0 et t_n . Les trajectoires sont par la suite ordonnées, de la première apparue, à la dernière, par exemple. Normalement il n'y a qu'une seule trajectoire qui apparaît à chaque instant t_k si le pas de discrétisation est petit, mais s'il y a deux apparitions au même point de temps (ce qui est possible en temps discret), on choisit de les ordonner selon leurs longueurs. Après l'observation complète du processus, on est en fine capable d'attribuer à chaque trajectoire un numéro.

5.2.1 Notations

Dans toute la suite, la notation $p(\cdot)$ correspond à la densité de probabilité (dans le cas conditionné ou non), que cette dernière soit au sens d'une probabilité de masse (cas discret) ou de la densité d'une loi continue.

Pour chaque trajectoire $i \in \llbracket 1, s \rrbracket$ on peut alors définir :

- $i_b \in \llbracket 0, n \rrbracket$ tel que la trajectoire i est apparue au temps t_{i_b}
- $i_d \in \llbracket 0, n \rrbracket$ tel que la trajectoire i a disparue au temps t_{i_d+1} , ie. t_{i_d} est son dernier instant de vie.

- $n_i := i_d - i_b + 1$ désigne le nombre d'instants de vie de la trajectoire. Sa durée de vie est donc $\Delta \times n_i$.

On note ensuite :

$$\mathbf{Z}^i := (z_{t_{i_b},i}, \dots, z_{t_{i_d},i}) \in \Lambda^{n_i}$$

les coordonnées de la trajectoire i .

On suppose que les particules peuvent suivre M types de mouvements différents, notés $1, \dots, M$ et qu'elles peuvent transiter d'un type de mouvement à un autre au cours de leur vie. L'ensemble des marques \mathbb{M} vaut donc $\{1, \dots, M\}$. On notera pour $k \in \llbracket i_b, i_d \rrbracket$, $m_{t_k,i}$ le type de mouvement que possède la particule i entre les temps t_{k-1} et t_k . (Ainsi, m_{t_0} sera défini mais n'aura aucune influence.)

On notera aussi :

$$\mathbf{M}^i = (m_{t_{i_b},i}, \dots, m_{t_{i_d},i}) \in \mathbb{M}^{n_i}$$

les états cachés de la trajectoire i .

On suppose ici que les mouvements des trajectoires et leurs changements de régimes sont indépendants, au sens de la définition 1 de l'annexe F. Intuitivement, cette définition implique que :

- le mouvement observé sur une trajectoire donnée dépend uniquement de la marque de la particule, indépendamment des autres trajectoires,
- lorsqu'une particule change de régime de mouvement, ce changement se fait indépendamment des autres particules.

Ainsi, pour chaque trajectoire \mathbf{Z}^i on a le schéma suivant :

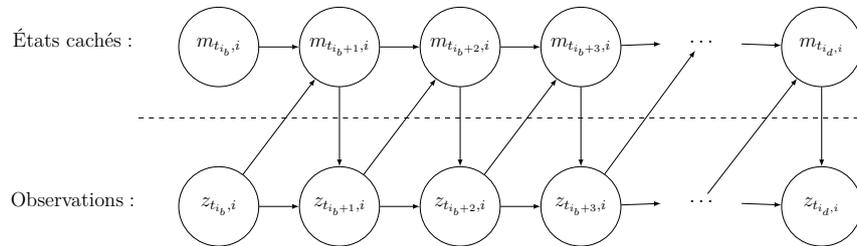


FIGURE 5.1 – Schéma des dépendances entre les observations et les états cachés (i.e. les types de mouvement) pour une trajectoire i .

Notons que des dépendances supplémentaires apparaissent dans le schéma de la figure 5.1 par rapport au cas standard des chaînes de Markov à états cachés.

Le système est entièrement décrit par les trois quantités suivantes :

- La loi de probabilité des changements d'états de chaque trajectoire est notée $a = \{a_{k,\ell}(z)\}$. Pour tout $i \in \llbracket 1, s \rrbracket$, pour tout $k \in \llbracket i_b + 1, i_d \rrbracket$, pour tout $(z, u, v) \in \Lambda \times \mathbb{M} \times \mathbb{M}$ on a :

$$a_{u,v}(z) = \mathbb{P} \left(m_{t_k,i} = v \mid z_{t_{k-1},i} = z, m_{t_{k-1},i} = u \right). \quad (5.1)$$

Nous faisons remarquer que cette loi ne dépend pas du temps t_k ni de la trajectoire i .

- Pour tout $i \in \llbracket 1, s \rrbracket$, pour tout $k \in \llbracket i_b + 1, i_d \rrbracket$, pour tout $(y, v) \in \Lambda \times \mathbb{M}$, la densité de $z_{t_k,i}$ sachant $z_{t_{k-1},i} = y$ et $m_{t_k,i} = v$ est donnée par la fonction

$$z \mapsto b^i(z|y, v). \quad (5.2)$$

- La loi initiale des couples $(z_{t_{i_b},i}, m_{t_{i_b},i})$ pour tout $i \in \llbracket 1, s \rrbracket$ est notée π :

$$\forall A \subset \Lambda, \quad \mathbb{P}(z_{t_{i_b},i} \in A, m_{t_{i_b},i} = k) = \int_A \pi(z, k) dz.$$

Le but de cette étude est à la fois d'estimer les états cachés $(m_{t_j,i})_{j \in \{i_b, \dots, i_d\}}$ pour tout $i \in \llbracket 1, s \rrbracket$ (et donc en déduire les instants de changement d'état), mais aussi de déterminer les caractéristiques des différents mouvements suivis (coefficient de diffusion, drift, etc). Pour cela, nous allons introduire dans le modèle un paramètre $\vartheta \in \mathbb{R}^k$ que l'on va chercher à estimer au sens du maximum de vraisemblance, grâce à un algorithme EM. Ce paramètre intervient dans la matrice a (car on cherche à déterminer la loi des changements de mouvement) mais aussi dans les densités b^i (qui caractérisent les mouvements suivis par les particules). En revanche, ce paramètre n'interviendra dans aucune autre caractéristique du processus BDMM. On rajoute une variable à ces quantités pour exprimer leur dépendance en ϑ : on écrit donc $b^i(z|y, v; \vartheta)$ au lieu de $b^i(z|y, v)$ et $a_{u,v}(z; \vartheta)$ au lieu de $a_{u,v}(z)$. De la même manière, on note $p(\cdot; \vartheta)$ au lieu de $p(\cdot)$.

Dans la partie 5.2.2, nous utilisons l'algorithme EM pour déterminer le paramètre qui maximise la vraisemblance. Dans la partie 5.2.3, nous détaillons comment

l'algorithme de Viterbi permet d'estimer les états cachés du problème, une fois que le paramètre ϑ optimal a été déterminé.

5.2.2 Algorithme EM pour estimer les paramètres qui maximisent la vraisemblance de la séquence observée

Il s'agit à présent de proposer une méthode permettant d'estimer les paramètres du modèle qui décrivent au mieux une séquence donnée d'observations. Formellement, le but est d'identifier le ϑ qui maximise la probabilité $p(\mathbf{Z}; \vartheta)$ avec :

$$\mathbf{Z} = (\mathbf{Z}^1, \dots, \mathbf{Z}^s).$$

Pour cela, on introduit la quantité :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) := \mathbb{E}_{\tilde{\vartheta}} (\ln(p(\mathbf{Z}, \mathbf{M}; \tilde{\vartheta}) | \mathbf{Z})),$$

avec :

$$\mathbf{M} = (\mathbf{M}^1, \dots, \mathbf{M}^s).$$

Proposition 1. *Si $Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) \geq Q(\vartheta, \vartheta, \mathbf{Z})$ alors $p(\mathbf{Z}, \tilde{\vartheta}) \geq p(\mathbf{Z}, \vartheta)$.*

Démonstration. On a :

$$\begin{aligned} Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= \mathbb{E}_{\tilde{\vartheta}} (\ln(p(\mathbf{Z}, \mathbf{M}; \tilde{\vartheta}) | \mathbf{Z})) \\ &= \sum_{\mathbf{m} \in \bigotimes_{i=1}^s \mathbb{M}^{n_i}} p(\mathbf{m} | \mathbf{Z}; \tilde{\vartheta}) \ln (p(\mathbf{Z}, \mathbf{m}; \tilde{\vartheta})). \end{aligned} \quad (5.3)$$

De plus,

$$\begin{aligned} \ln \left(\frac{p(\mathbf{Z}; \tilde{\vartheta})}{p(\mathbf{Z}; \vartheta)} \right) &= \ln \left(\frac{\sum_{\mathbf{m} \in \bigotimes_{i=1}^s \mathbb{M}^{n_i}} p(\mathbf{Z}, \mathbf{m}; \tilde{\vartheta})}{p(\mathbf{Z}; \vartheta)} \right) \\ &= \ln \left(\frac{1}{p(\mathbf{Z}; \vartheta)} \sum_{\mathbf{m} \in \bigotimes_{i=1}^s \mathbb{M}^{n_i}} p(\mathbf{Z}, \mathbf{m}; \vartheta) \frac{p(\mathbf{Z}, \mathbf{m}; \tilde{\vartheta})}{p(\mathbf{Z}, \mathbf{m}; \vartheta)} \right). \end{aligned}$$

En raison de la concavité de la fonction logarithme, et que

$$\frac{1}{p(\mathbf{Z}; \vartheta)} \sum_{\mathbf{m} \in \bigotimes_{i=1}^s \mathbb{M}^{n_i}} p(\mathbf{Z}, \mathbf{m}; \vartheta) = 1,$$

on a :

$$\begin{aligned} \ln \left(\frac{p(\mathbf{Z}; \tilde{\vartheta})}{p(\mathbf{Z}; \vartheta)} \right) &\geq \frac{1}{p(\mathbf{Z}; \vartheta)} \sum_{\mathbf{m} \in \bigotimes_{i=1}^s \mathbb{M}^{n_i}} p(\mathbf{Z}, \mathbf{m}; \vartheta) \ln \left(\frac{p(\mathbf{Z}, \mathbf{m}; \tilde{\vartheta})}{p(\mathbf{Z}, \mathbf{m}; \vartheta)} \right) \\ &= \sum_{\mathbf{m} \in \bigotimes_{i=1}^s \mathbb{M}^{n_i}} \frac{p(\mathbf{Z}, \mathbf{m}; \vartheta)}{p(\mathbf{Z}; \vartheta)} \ln \left(\frac{p(\mathbf{Z}, \mathbf{m}; \tilde{\vartheta})}{p(\mathbf{Z}, \mathbf{m}; \vartheta)} \right) \\ &= \sum_{\mathbf{m} \in \bigotimes_{i=1}^s \mathbb{M}^{n_i}} p(\mathbf{m}|\mathbf{Z}; \vartheta) \ln \left(\frac{p(\mathbf{Z}, \mathbf{m}; \tilde{\vartheta})}{p(\mathbf{Z}, \mathbf{m}; \vartheta)} \right) \\ &= Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) - Q(\vartheta, \vartheta, \mathbf{Z}). \end{aligned}$$

□

On peut donc se servir de la fonction Q pour approcher le maximum de vraisemblance, selon le principe adopté dans l’algorithme EM.

5.2.2.1 Algorithme EM

L’algorithme EM est un algorithme itératif qui alterne entre deux étapes : une étape dite “E” (Expectation) qui consiste à calculer l’espérance conditionnelle de la log-vraisemblance des données observées et des états cachés sachant les données observées et les paramètres inconnus estimés à l’étape précédente ; une étape dite “M” (Maximization) qui consiste à mettre à jour les paramètres en maximisant l’espérance conditionnelle. Ainsi, à partir de la valeur initiale $\vartheta^{[0]}$, l’itération k est définie comme suit :

- **Étape E** : calcul de l’espérance de la log-vraisemblance des données observées et des états cachés sachant les données observées et les paramètres inconnus actuels $\vartheta^{[k]}$:

$$Q(\tilde{\vartheta}, \vartheta^{[k]}, \mathbf{Z}).$$

— **Étape M** : mise à jour des paramètres en maximisant $Q(\tilde{\vartheta}, \vartheta^{[k]}, \mathbf{Z})$ sur $\tilde{\vartheta}$:

$$\vartheta^{[k+1]} = \arg \max_{\tilde{\vartheta}} Q(\tilde{\vartheta}, \vartheta^{[k]}, \mathbf{Z}).$$

L'algorithme s'arrête lorsque, pour une itération $[s]$, on a :

$$|Q(\vartheta^{[k]}, \vartheta^{[k-1]}, \mathbf{Z}) - Q(\vartheta^{[k-1]}, \vartheta^{[k-1]}, \mathbf{Z})| < \eta,$$

pour un seuil fixé de tolérance η .

L'algorithme EM est monotone croissant (c'est-à-dire qu'il augmente la log-vraisemblance des données observées à chaque itération) et converge donc vers un maximum local de la fonction objective. En outre, cet algorithme est déterministe, ce qui signifie que le point de convergence ne dépend que du point de départ $\vartheta^{[0]}$. Par conséquent, cet algorithme doit être exécuté avec différents points de départ (généralement échantillonnés de manière aléatoire) pour espérer atteindre le maximum global de vraisemblance.

Afin de mettre en place cet algorithme EM, il faut donc calculer l'argument du maximum de $Q(\tilde{\vartheta}, \vartheta, \mathbf{Z})$. Pour cela, nous introduisons de nouvelles variables intermédiaires.

5.2.2.2 Algorithme Forward-Backward pour calculer Q

Dans toute la suite, on note pour tout $i \in \llbracket 1, s \rrbracket$ et pour tout $j, k \in \llbracket i_b, i_d \rrbracket$ avec $j \leq k$, $\mathbf{Z}_{t_j:t_k}^i = (z_{t_j,i}, \dots, z_{t_k,i})$.

Pour chaque trajectoire i , nous introduisons quatre nouvelles variables : α^i , β^i , γ^i , ξ^i puis nous donnons une formule de récurrence pour le calcul de chacune d'entre elles. Ces quatre variables nous seront nécessaires pour déterminer la valeur de la fonction Q , via l'expression donnée dans la proposition 6.

Pour tout $i \in \llbracket 1, s \rrbracket$ et pour tout $k \in \llbracket i_b, i_d \rrbracket$, on introduit la variable "forward",

qui est la loi jointe de $\mathbf{Z}_{t_b:t_k}^i$ et $m_{t_k,i}$:

$$\alpha_k^i(j; \vartheta) = p(\mathbf{Z}_{t_b:t_k}^i, m_{t_k,i} = j; \vartheta).$$

Le coefficient $\alpha_k^i(j; \vartheta)$ correspond à la probabilité que les $(k - i_b)$ premières observations de la trajectoire i , $\mathbf{Z}_{t_b:t_k}^i$, aient lieu et que le $(k - i_b)$ -ième état caché soit le type de mouvement j .

Proposition 2 (Algorithme “forward” de Baum). *La suite des $(\alpha_k^i(j; \vartheta))_{k \in \llbracket i_b, i_d \rrbracket}$ s’obtient récursivement de la façon suivante :*

$$\begin{aligned} \alpha_{i_b}^i(j; \vartheta) &= \pi(z_{t_b,i}, j), \\ \text{et } \forall k \in \llbracket i_b + 1, i_d \rrbracket, \quad \alpha_k^i(j; \vartheta) &= \left(\sum_{\ell=1}^M \alpha_{k-1}^i(\ell; \vartheta) a_{\ell,j}(z_{t_{k-1},i}; \vartheta) \right) b^i(z_{t_k,i} | z_{t_{k-1},i}, j; \vartheta). \end{aligned}$$

Démonstration. Soit $j \in \llbracket 1, M \rrbracket$. On a :

$$\alpha_{i_b}^i(j; \vartheta) = p(\mathbf{Z}_{t_b:t_b}^i, m_{t_b,i} = j; \vartheta) = p(z_{t_b,i}, m_{t_b,i} = j; \vartheta) = \pi(z_{t_b,i}, j).$$

Soit $k > i_b$,

$$\begin{aligned} \alpha_k^i(j; \vartheta) &= p(\mathbf{Z}_{t_b:t_k}^i, m_{t_k,i} = j; \vartheta) \\ &= \sum_{\ell=1}^M p(\mathbf{Z}_{t_b:t_k}^i, m_{t_k,i} = j, m_{t_{k-1},i} = \ell; \vartheta) \\ &= \sum_{\ell=1}^M p(z_{t_k,i}, \mathbf{Z}_{t_b:t_{k-1}}^i, m_{t_k,i} = j, m_{t_{k-1},i} = \ell; \vartheta) \end{aligned}$$

donc :

$$\begin{aligned} \alpha_k^i(j; \vartheta) &= \sum_{\ell=1}^M p(z_{t_k,i}, m_{t_k,i} = j | \mathbf{Z}_{t_b:t_{k-1}}^i, m_{t_{k-1},i} = \ell; \vartheta) p(\mathbf{Z}_{t_b:t_{k-1}}^i, m_{t_{k-1},i} = \ell; \vartheta) \\ &= \sum_{\ell=1}^M p(z_{t_k,i} | m_{t_k,i} = j, \mathbf{Z}_{t_b:t_{k-1}}^i, m_{t_{k-1},i} = \ell; \vartheta) \\ &\quad \times p(m_{t_k,i} = j | \mathbf{Z}_{t_b:t_{k-1}}^i, m_{t_{k-1},i} = \ell; \vartheta) \alpha_{k-1}^i(\ell; \vartheta) \\ &= \sum_{\ell=1}^M b^i(z_{t_k,i} | z_{t_{k-1},i}, j; \vartheta) a_{\ell,j}(z_{t_{k-1},i}; \vartheta) \alpha_{k-1}^i(\ell; \vartheta). \end{aligned}$$

□

Pour tout $i \in \llbracket 1, s \rrbracket$ et pour tout $k \in \llbracket i_b, i_d \rrbracket$, on introduit la variable “backward” (rétroactive) pour chaque trajectoire \mathbf{Z}^i :

$$\beta_k^i(j; \vartheta) = p(\mathbf{Z}_{t_{k+1}:t_{i_d}}^i | m_{t_k,i} = j, z_{t_k,i}; \vartheta),$$

avec la convention : $\beta_{i_d}^i(j; \vartheta) = 1$ pour tout $j \in \llbracket 1, M \rrbracket$.

Le coefficient $\beta_k^i(j; \vartheta)$ correspond à la probabilité d’observer la séquence partielle ultérieure $\mathbf{Z}_{t_{k+1}:i_d}^i$, sachant que la trajectoire i de paramètre ϑ suit le type de mouvement j à l’instant t_k et connaissant le dernier point de la trajectoire antérieure $z_{t_k,i}$.

Proposition 3 (Algorithme “backward” de Baum). *La suite des $(\beta_k^i(j; \vartheta))_{k \in \llbracket i_b, i_d \rrbracket}$ s’obtient récursivement de la façon suivante :*

$$\forall k \in \llbracket i_b, i_d - 1 \rrbracket, \quad \beta_k^i(j; \vartheta) = \sum_{\ell=1}^M \beta_{k+1}^i(\ell; \vartheta) a_{j,\ell}(z_{t_k,i}; \vartheta) b^i(z_{t_{k+1},i} | z_{t_k,i}, \ell; \vartheta).$$

Démonstration. Soit $k \in \llbracket i_b, i_d - 1 \rrbracket$, on a :

$$\begin{aligned} \beta_k^i(j; \vartheta) &= p(\mathbf{Z}_{t_{k+1}:t_{i_d}}^i | m_{t_k,i} = j, z_{t_k,i}; \vartheta) \\ &= \sum_{\ell=1}^M p(\mathbf{Z}_{t_{k+1}:t_{i_d}}^i, m_{t_{k+1},i} = \ell | m_{t_k,i} = j, z_{t_k,i}; \vartheta) \\ &= \sum_{\ell=1}^M p(\mathbf{Z}_{t_{k+1}:t_{i_d}}^i | m_{t_{k+1},i} = \ell, m_{t_k,i} = j, z_{t_k,i}; \vartheta) p(m_{t_{k+1},i} = \ell | m_{t_k,i} = j, z_{t_k,i}; \vartheta) \\ &= \sum_{\ell=1}^M p(\mathbf{Z}_{t_{k+2}:t_{i_d}}^i, z_{t_{k+1},i} | m_{t_{k+1},i} = \ell, m_{t_k,i} = j, z_{t_k,i}; \vartheta) a_{j,\ell}(z_{t_k,i}; \vartheta) \\ &= \sum_{\ell=1}^M p(\mathbf{Z}_{t_{k+2}:t_{i_d}}^i | z_{t_{k+1},i}, m_{t_{k+1},i} = \ell, m_{t_k,i} = j, z_{t_k,i}; \vartheta) \\ &\quad \times p(z_{t_{k+1},i} | m_{t_{k+1},i} = \ell, m_{t_k,i} = j, z_{t_k,i}; \vartheta) a_{j,\ell}(z_{t_k,i}; \vartheta) \\ &= \sum_{\ell=1}^M p(\mathbf{Z}_{t_{k+2}:t_{i_d}}^i | z_{t_{k+1},i}, m_{t_{k+1},i} = \ell; \vartheta) p(z_{t_{k+1},i} | m_{t_{k+1},i} = \ell, z_{t_k,i}; \vartheta) a_{j,\ell}(z_{t_k,i}; \vartheta) \\ &= \sum_{\ell=1}^M \beta_{k+1}^i(\ell; \vartheta) b^i(z_{t_{k+1},i} | z_{t_k,i}, \ell; \vartheta) a_{j,\ell}(z_{t_k,i}; \vartheta). \end{aligned}$$

□

On introduit :

$$\gamma_k^i(j; \vartheta) = p(m_{t_k, i} = j | \mathbf{Z}^i; \vartheta) \quad (5.4)$$

et

$$\xi_k^i(j, \ell; \vartheta) = p(m_{t_k, i} = j, m_{t_{k+1}, i} = \ell | \mathbf{Z}^i; \vartheta). \quad (5.5)$$

Le coefficient $\xi_k^i(j, \ell; \vartheta)$ représente la probabilité selon le modèle de paramètres ϑ et la séquence d'observations \mathbf{Z}^i de passer du type de mouvement j au moment t_k au type de mouvement ℓ au moment t_{k+1} . Le coefficient $\gamma_k^i(j; \vartheta)$, quant à lui, correspond à la probabilité selon le modèle de paramètres ϑ et la séquence d'observations \mathbf{Z}^i d'être dans le type de mouvement j à l'instant t_k .

Proposition 4 (Algorithme Forward-Backward). *Pour $k \in \llbracket i_b, i_d \rrbracket$, on a :*

$$\gamma_k^i(j; \vartheta) = \frac{\alpha_k^i(j; \vartheta) \beta_k^i(j; \vartheta)}{\sum_{j=1}^M \beta_k^i(j; \vartheta) \alpha_k^i(j; \vartheta)}.$$

Démonstration.

$$\gamma_k^i(j; \vartheta) = p(m_{t_k, i} = j | \mathbf{Z}^i; \vartheta) = \frac{p(m_{t_k, i} = j, \mathbf{Z}^i; \vartheta)}{p(\mathbf{Z}^i; \vartheta)}.$$

Le dénominateur se calcule ainsi :

$$\begin{aligned} p(\mathbf{Z}^i; \vartheta) &= \sum_{j=1}^M p(\mathbf{Z}^i, m_{t_k, i} = j; \vartheta) \\ &= \sum_{j=1}^M p(\mathbf{Z}_{t_{k+1}:t_{i_d}}^i | m_{t_k, i} = j, \mathbf{Z}_{t_{i_b}:t_k}^i; \vartheta) p(m_{t_k, i} = j, \mathbf{Z}_{t_{i_b}:t_k}^i; \vartheta) \\ &= \sum_{j=1}^M p(\mathbf{Z}_{t_{k+1}:t_{i_d}}^i | m_{t_k, i} = j, z_{t_k, i}; \vartheta) \alpha_k^i(j; \vartheta) \\ &= \sum_{j=1}^M \beta_k^i(j; \vartheta) \alpha_k^i(j; \vartheta). \end{aligned}$$

Le numérateur vaut :

$$\begin{aligned}
 p(m_{t_k,i} = j, \mathbf{Z}^i; \vartheta) &= p(\mathbf{Z}_{t_{k+1}:t_{i_d}}^i | m_{t_k,i} = j, \mathbf{Z}_{t_{i_b}:t_k}^i; \vartheta) p(m_{t_k,i} = j, \mathbf{Z}_{t_{i_b}:t_k}^i; \vartheta) \\
 &= p(\mathbf{Z}_{t_{k+1}:t_{i_d}}^i | m_{t_k,i} = j, z_{t_k,i}; \vartheta) \alpha_k^i(j; \vartheta) \\
 &= \beta_k^i(j; \vartheta) \alpha_k^i(j; \vartheta).
 \end{aligned}$$

□

Proposition 5. Pour $k \in \llbracket i_b, i_d - 1 \rrbracket$, on a :

$$\xi_k^i(j, \ell; \vartheta) = \frac{\alpha_k^i(j; \vartheta) a_{j,\ell}(\vartheta) b^i(z_{t_{k+1},i}, z_{t_k,i}, \ell; \vartheta) \beta_{t_{k+1}}^i(\ell; \vartheta)}{\sum_{j=1}^M \beta_k^i(j; \vartheta) \alpha_k^i(j; \vartheta)}.$$

Démonstration.

$$\xi_k^i(j, \ell; \vartheta) = p(m_{t_k,i} = j, m_{t_{k+1},i} = \ell | \mathbf{Z}^i; \vartheta) = \frac{p(m_{t_k,i} = j, m_{t_{k+1},i} = \ell, \mathbf{Z}^i; \vartheta)}{p(\mathbf{Z}^i; \vartheta)}.$$

Le dénominateur a déjà été calculé précédemment (preuve de la proposition 4). Le numérateur vaut :

$$\begin{aligned}
 p(m_{t_k,i} = j, m_{t_{k+1},i} = \ell, \mathbf{Z}^i; \vartheta) &= p(\mathbf{Z}_{t_{k+1}:t_{i_d}}^i | m_{t_k,i} = j, m_{t_{k+1},i} = \ell, \mathbf{Z}_{t_{i_b}:t_k}^i; \vartheta) \\
 &\quad \times p(m_{t_{k+1},i} = \ell | m_{t_k,i} = j, \mathbf{Z}_{t_{i_b}:t_k}^i; \vartheta) p(m_{t_k,i} = j, \mathbf{Z}_{t_{i_b}:t_k}^i; \vartheta)
 \end{aligned}$$

donc :

$$\begin{aligned}
 p(m_{t_k,i} = j, m_{t_{k+1},i} = \ell, \mathbf{Z}^i; \vartheta) &= p(\mathbf{Z}_{t_{k+2}:t_{i_d}}^i | m_{t_k,i} = j, m_{t_{k+1},i} = \ell, \mathbf{Z}_{t_{i_b}:t_{k+1}}^i; \vartheta) \\
 &\quad \times p(z_{t_{k+1},i} | m_{t_k,i} = j, m_{t_{k+1},i} = \ell, \mathbf{Z}_{t_{i_b}:t_k}^i; \vartheta) a_{j,\ell}(z_{t_k,i}; \vartheta) \alpha_k(j; \vartheta) \\
 &= p(\mathbf{Z}_{t_{k+2}:t_{i_d}}^i | m_{t_{k+1},i} = \ell, z_{t_{k+1},i}; \vartheta) \\
 &\quad \times p(z_{t_{k+1},i} | m_{t_{k+1},i} = \ell, z_{t_k,i}; \vartheta) a_{j,\ell}(z_{t_k,i}; \vartheta) \alpha_k(j; \vartheta) \\
 &= \beta_{k+1}(k; \vartheta) b^i(z_{t_{k+1},i} | z_{t_k,i}, k; \vartheta) a_{j,\ell}(z_{t_k,i}; \vartheta) \alpha_k(j; \vartheta).
 \end{aligned}$$

□

On introduit trois conditions :

- (C1) L'intensité de saut α est bornée.
- (C2) Le paramètre d'intérêt ϑ n'intervient que dans τ , k_τ et L .
- (C3) Les mouvements des trajectoires et leurs changements de régime sont indépendants au sens de la définition 1 de l'annexe F .

La proposition suivante permet d'exprimer la quantité $Q(\tilde{\vartheta}, \vartheta, \mathbf{Z})$ comme une somme trajectoire par trajectoire, en fonction des variables précédemment introduites. Sa preuve constitue le coeur de l'annexe F.

Proposition 6. *Si les conditions (C1), (C2) et (C3) sont vérifiées, alors la fonction Q s'écrit :*

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \bar{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) \left(1 + \underset{\Delta \rightarrow 0}{o_{\mathbb{P}}}(1)\right) + \hat{C}(\vartheta, \mathbf{Z}, \Delta).$$

avec :

$$\begin{aligned} \bar{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = & \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \sum_{u=1}^M \sum_{v=1}^M \xi_k^i(u, v; \vartheta) \ln(a_{u,v}(z_{t_{k-1}, i}; \tilde{\vartheta})) \\ & + \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \sum_{v=1}^M \gamma_k^i(v; \vartheta) \ln\left(b^i(z_{t_k, i} | z_{t_{k-1}, i}, v; \tilde{\vartheta})\right) \end{aligned}$$

où les expressions de γ_k^i et ξ_k^i sont respectivement données par les propositions 4 et 5.

5.2.3 Algorithme de Viterbi pour estimer les états cachés ayant produit la séquence observée

Dans cette section, le but est de déterminer la partie cachée du modèle, c'est-à-dire de déterminer les suites des états cachés $(m_{t_k, i})_{k \in \llbracket i_b, i_d \rrbracket}$ pour tout $i \in \llbracket 1, s \rrbracket$ qui sont respectivement les plus probables d'avoir engendré les suites des observations \mathbf{Z}^i avec le paramètre ϑ optimal déterminé dans la section précédente. Pour cela, on recourt à une adaptation de l'algorithme de Viterbi, déjà bien connu pour les modèles de Markov à états cachés standards.

L'objectif est de calculer :

$$\arg \max_{\mathbf{m} \in \mathbb{M}^{n_i}} p(\mathbf{M}_{t_{i_b}:t_{i_d}}^i = \mathbf{m} | \mathbf{Z}^i; \vartheta). \quad (5.6)$$

On fait remarquer que :

$$\begin{aligned} \arg \max_{\mathbf{m} \in \mathbb{M}^{n_i}} p(\mathbf{M}_{t_{i_b}:t_{i_d}}^i = \mathbf{m} | \mathbf{Z}^i; \vartheta) &= \arg \max_{\mathbf{m} \in \mathbb{M}^{i_d - i_b + 1}} \frac{p(\mathbf{M}_{t_{i_b}:t_{i_d}}^i = \mathbf{m}, \mathbf{Z}^i; \vartheta)}{p(\mathbf{Z}^i; \vartheta)} \\ &= \arg \max_{\mathbf{m} \in \mathbb{M}^{i_d - i_b + 1}} p(\mathbf{M}_{t_{i_b}:t_{i_d}}^i = \mathbf{m}, \mathbf{Z}^i; \vartheta). \end{aligned}$$

Ainsi, il est équivalent dans (5.6) de chercher l'argument du maximum de la loi conditionnelle ou de la loi jointe. Pour répondre à ce problème, nous allons utiliser un algorithme de programmation dynamique : l'idée générale est de prendre une décision optimale à chaque étape de la trajectoire pour chaque état possible.

On introduit alors pour tout $k \in \llbracket i_b, i_d \rrbracket$:

$$\delta_k^i(j; \vartheta) := \max_{\mathbf{m} \in \mathbb{M}^{k - i_b}} p(\mathbf{M}_{t_{i_b}:t_{k-1}}^i = \mathbf{m}, m_{t_k, i} = j, \mathbf{Z}_{t_{i_b}:t_k}^i; \vartheta)$$

avec la convention $\delta_{i_b}^i(j; \vartheta) = \pi(z_{t_{i_b}, i}, j)$.

Le coefficient $\delta_k^i(j; \vartheta)$ représente la probabilité que le $(k - i_b)$ -ième état caché soit le type de mouvement j après avoir observé les $(k - i_b)$ premiers points de la trajectoire i et être passé par la suite des $(k - i_b)$ premiers états cachés la plus probable sachant le paramètre ϑ .

Proposition 7 (Algorithme de programmation dynamique). *La suite $(\delta_k^i(j; \vartheta))_{k \in \llbracket i_b + 1, i_d \rrbracket}$ s'obtient récursivement de la façon suivante :*

$$\forall k \in \llbracket i_b + 1, i_d \rrbracket, \quad \delta_k^i(j; \vartheta) = \max_{1 \leq \ell \leq M} \left(\delta_{k-1}^i(\ell; \vartheta) a_{\ell, j}^i(\vartheta) \right) b^i(z_{t_k, i} | z_{t_{k-1}, i}, j; \vartheta).$$

Démonstration. Soit $k \in \llbracket i_b, i_d \rrbracket$. On a :

$$\begin{aligned} \delta_k^i(j; \vartheta) &= \max_{\mathbf{m}_{i_b:k-1} \in \mathbb{M}^{k-i_b}} p(\mathbf{M}_{t_{i_b}:t_{k-1}}^i = \mathbf{m}_{i_b:k-1}, m_{t_k,i} = j, \mathbf{Z}_{t_{i_b}:t_k}^i; \vartheta) \\ &= \max_{\mathbf{m}_{i_b:k-1} \in \mathbb{M}^{k-i_b}} p(z_{t_k,i} | \mathbf{Z}_{t_{i_b}:t_{k-1}}^i, \mathbf{M}_{t_{i_b}:t_{k-1}}^i = \mathbf{m}_{i_b:k-1}, m_{t_k,i} = j; \vartheta) \\ &\quad \times p(m_{t_k,i} = j | \mathbf{M}_{t_{i_b}:t_{k-1}}^i = \mathbf{m}_{i_b:k-1}, \mathbf{Z}_{t_{i_b}:t_{k-1}}^i; \vartheta) \\ &\quad \times p(\mathbf{Z}_{t_{i_b}:t_{k-1}}^i, \mathbf{M}_{t_{i_b}:t_{k-2}}^i = \mathbf{m}_{i_b:k-2}, m_{t_{k-1},i} = x_{k-1}; \vartheta) \end{aligned}$$

donc :

$$\begin{aligned} \delta_k^i(j; \vartheta) &= \max_{\mathbf{m}_{i_b:k-1} \in \mathbb{M}^{k-i_b}} p(z_{t_k,i} | z_{t_{k-1},i}, m_{t_k,i} = j; \vartheta) p(m_{t_k,i} = j | m_{t_{k-1},i} = m_{k-1}, z_{t_{k-1},i}; \vartheta) \\ &\quad \times p(\mathbf{Z}_{t_{i_b}:t_{k-1}}^i, \mathbf{M}_{t_{i_b}:t_{k-2}}^i = \mathbf{m}_{i_b:k-2}, m_{t_{k-1},i} = m_{k-1}; \vartheta) \\ &= \max_{\mathbf{m}_{i_b:k-1} \in \mathbb{M}^{k-i_b}} b^i(z_{t_k,i} | z_{t_{k-1},i}, j; \vartheta) a_{m_{k-1},j}(z_{t_{k-1},i}; \vartheta) \\ &\quad \times p(\mathbf{Z}_{t_{i_b}:t_{k-1}}^i, \mathbf{M}_{t_{i_b}:t_{k-2}}^i = \mathbf{m}_{i_b:k-2}, m_{t_{k-1},i} = m_{k-1}; \vartheta) \\ &= \max_{x_{k-1} \in \mathbb{M}} b^i(z_{t_k,i} | z_{t_{k-1},i}, j; \vartheta) a_{m_{k-1},j}(z_{t_{k-1},i}; \vartheta) \\ &\quad \times \max_{\mathbf{m}_{i_b:k-2} \in \mathbb{M}^{k-i_b-1}} p(\mathbf{Z}_{t_{i_b}:t_{k-1}}^i, \mathbf{M}_{t_{i_b}:t_{k-2}}^i = \mathbf{m}_{i_b:k-2}, m_{t_{k-1},i} = m_{k-1}; \vartheta) \\ &= \max_{m_{k-1} \in \mathbb{M}} b^i(z_{t_k,i} | z_{t_{k-1},i}, j; \vartheta) a_{m_{k-1},j}(z_{t_{k-1},i}; \vartheta) \delta_{k-1}^i(m_{k-1}; \vartheta). \end{aligned}$$

□

Nous pouvons obtenir l'estimateur $\widehat{\mathbf{M}}^i = (\widehat{m}_{t_{i_b},i}, \dots, \widehat{m}_{t_{i_d},i})$ qui est la suite la plus probable des états cachés de la trajectoire i définis comme :

$$\widehat{\mathbf{M}}^i = \arg \max_{\mathbf{m} \in \mathbb{M}^{n_i}} p(\mathbf{M}^i = \mathbf{m} | \mathbf{Z}^i; \vartheta)$$

grâce à l'algorithme suivant.

Proposition 8 (Algorithme de Viterbi). *L'estimateur des états cachés*

$$\widehat{\mathbf{M}}^i = (\widehat{m}_{t_{i_b},i}, \dots, \widehat{m}_{t_{i_d},i})$$

s'obtient récursivement de la façon suivante :

$$\widehat{m}_{t_{i_d}, i} = \arg \max_{j \in \mathbb{M}} \delta_{i_d}^i(j; \vartheta),$$

et pour tout $k = i_d - 1, \dots, 1$,

$$\widehat{m}_{t_k, i} = \arg \max_{j \in \mathbb{M}} \delta_k^i(j; \vartheta) a_{j, \widehat{m}_{t_{k+1}, i}}(z_{t_k, i}; \vartheta).$$

Démonstration. On remarque que :

$$\begin{aligned} & \max_{\mathbf{m}_{i_b:i_d} \in \mathbb{M}^{n_i}} p(\mathbf{M}_{t_{i_b}:t_{i_d}}^i = \mathbf{m}_{i_b:i_d}, \mathbf{Z}^i; \vartheta) \\ &= \max_{m_{i_d} \in \mathbb{M}} \max_{\mathbf{m}_{i_b:i_{d-1}} \in \mathbb{M}^{i_d-i_b-2}} p(\mathbf{M}_{t_{i_b}:t_{i_d}}^i = \mathbf{m}_{i_b:i_{d-1}}, m_{t_{i_d}, i} = m_{i_d}, \mathbf{Z}^i; \vartheta) \\ &= \max_{m_{i_d} \in \mathbb{M}} \delta_{i_d}^i(m_{i_d}; \vartheta), \end{aligned}$$

donc le $(i_d - i_b - 2)$ -ième état caché le plus probable est :

$$\widehat{m}_{t_{i_d}, i} = \arg \max_{j \in \mathbb{M}} \delta_{i_d}^i(j; \vartheta).$$

Parmi toutes les suites d'états cachés qui se trouvent dans l'état x' à un instant k , celle de plus grande probabilité se trouve à l'instant $k - 1$ dans l'état caché :

$$\arg \max_{x \in \mathbb{M}} \delta_{k-1}^i(x; \vartheta) a_{x, x'}(z_{t_{k-1}, i}; \vartheta).$$

Ainsi, de proche en proche, on a :

$$\widehat{m}_{t_{i_d-1}, i} = \arg \max_{j \in \mathbb{M}} \delta_{i_d-1}^i(j; \vartheta) a_{j, \widehat{m}_{t_{i_d}, i}}(z_{t_{i_d-1}, i}; \vartheta),$$

puis

$$\widehat{m}_{t_{i_d-2}, i} = \arg \max_{j \in \mathbb{M}} \delta_{i_d-2}^i(j; \vartheta) a_{j, \widehat{m}_{t_{i_d-1}, i}}(z_{t_{i_d-2}, i}; \vartheta),$$

et on continue ainsi jusqu'à $k = i_b$. Pour $k \in \llbracket i_b, i_d \rrbracket$, on a :

$$\widehat{m}_{t_k, i} = \arg \max_{j \in \mathbb{M}} \delta_k^i(j; \vartheta) a_{j, \widehat{m}_{t_{k+1}, i}}(z_{t_k, i}; \vartheta).$$

□

5.3 Étude théorique d'un exemple dans le cas de trois mouvements

Dans cette section, nous présentons un exemple d'application de la théorie présentée dans la section précédente. Les caractéristiques sont judicieusement choisies afin d'approcher la dynamique des données réelles que nous étudierons dans le chapitre 6, et déjà étudiées dans les chapitres 2 et 4.

Dans la sous-section 5.3.1, nous présentons l'exemple et ses caractéristiques, puis nous en déduisons les expressions des fonctions a et b^i correspondantes. La section F.4 de l'annexe F montre que cet exemple satisfait les conditions nécessaires pour appliquer la proposition 6, ce qui nous permettra d'obtenir une expression de la fonction Q . Dans la sous-section 5.3.2, nous nous servons de cette expression pour en déduire une expression de l'argument du maximum de la fonction Q , qui est explicite dans cet exemple.

5.3.1 Présentation de l'exemple

Dans cet exemple, on considère un processus BDMM dont l'espace de vie des individus Λ est inclu dans \mathbb{R}^2 , et qui respecte les hypothèses suivantes (en gardant les notations du chapitre 3) :

- La marque de chaque particule identifie son régime de mouvement à chaque instant. Ce régime peut être de trois types différents (qui seront décrits ci-dessous). On note l'ensemble \mathbb{M} des marques ainsi :

$$\mathbb{M} = \{1, 2, 3\}.$$

- Chaque particule possède un taux de mutation constant égal à τ :

$$\tau(\mathbf{x}) = \tau \times n(\mathbf{x}).$$

- Le noyau de mutation est celui de l'exemple 3 du chapitre 3, c'est-à-dire que

pour $(\mathbf{x}, \mathbf{y}) \in E^2$, on a :

$$k_\tau(\mathbf{x}, \mathbf{y}) = \begin{cases} \frac{1}{n(\mathbf{x})} p_{m_i, m} & \text{si } \exists x_i = (z_i, m_i) \in \mathbf{x}, \exists m \in \mathbb{M}, \mathbf{y} = (\mathbf{x} \setminus (z_i, m_i)) \cup (z_i, m), \\ 0 & \text{sinon,} \end{cases}$$

où $(p_{i,j})_{(i,j) \in \{1,2,3\}^2}$ est une matrice de transition de la forme suivante :

$$\begin{pmatrix} 0 & p_b & 1 - p_b \\ p_r & 0 & 1 - p_r \\ p_v & 1 - p_v & 0 \end{pmatrix}.$$

Le coefficient p_b (resp. p_r, p_v) apparaît comme la probabilité de passer de l'état **1** à **2** (resp. **2** à **1**, **3** à **1**), sachant que la particule qui mute est dans l'état **1** (resp. **2**, **3**).

- Les particules changent de position chacune selon sa marque et indépendamment des autres.
- L'intensité de mort δ est quelconque mais bornée.
- L'intensité de naissance β est quelconque, bornée et vérifie :

$$\exists n^* \in \mathbb{N}, \forall \mathbf{x} \in E, \quad n(\mathbf{x}) \geq n^* \implies \beta(\mathbf{x}) = 0.$$

- Les noyaux de naissance et de mort sont quelconques.
- Le paramètre d'intérêt ϑ n'intervient que dans l'intensité de mutation τ , dans le noyau de mutation k_τ et dans les paramètres des mouvements que suivent les particules entre les sauts.

À présent, nous définissons les trois types de mouvement, notés **1**, **2**, **3** ainsi :

- **Type 1** : mouvement brownien. Cela signifie que si $m_{t_k, i} = \mathbf{1}$, la particule i est, durant le pas de discrétisation $[t_{k-1}, t_k]$, solution de l'équation :

$$dX_t = \sigma_{i,1} dB_t,$$

où B_t est un mouvement brownien standard et où le coefficient de diffusion $\sigma_{i,1} > 0$ est un paramètre inconnu, inclu dans le vecteur d'inconnues ϑ .

- **Type 2** : mouvement superdiffusif qui est ici un mouvement brownien avec dérive. Cela signifie que si $m_{t_k, i} = 2$, la particule i est, durant le pas de discrétisation $[t_{k-1}, t_k]$, solution de l'équation :

$$dX_t = \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} dt + \sigma_{i,2} dB_t,$$

où le coefficient de diffusion $\sigma_{i,2} > 0$, la norme du drift $\rho_i > 0$ et la direction du drift $\eta_i \in [0, 2\pi[$ sont des paramètres inconnus, inclus dans le vecteur d'inconnues ϑ .

- **Type 3** : mouvement confiné qui est ici solution d'un processus d'Ornstein-Uhlenbeck. Cela signifie que si $m_{t_k, i} = 3$, la particule i est, durant le pas de discrétisation $[t_{k-1}, t_k]$, solution de l'équation :

$$dX_t = -\lambda_i(X_t - \mu_i)dt + \sigma_{i,3}dB_t,$$

où le coefficient de diffusion $\sigma_{i,3} > 0$, la force de rappel $\lambda_i \geq 0$ et le point d'équilibre $\mu_i = \begin{pmatrix} \mu_i(0) \\ \mu_i(1) \end{pmatrix} \in \Lambda \subset \mathbb{R}^2$ sont des paramètres inconnus, inclus dans le vecteur d'inconnues ϑ .

Ainsi, pour une trajectoire i , il y a 8 paramètres réels inconnus (μ_i intégrant les deux paramètres $\mu_i(0)$ et $\mu_i(1)$) :

$$(\sigma_{i,1}^2, \rho_i, \eta_i, \sigma_{i,2}^2, \lambda_i, \mu_i, \sigma_{i,3}^2) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times [0, 2\pi[\times \mathbb{R}_+^* \times \mathbb{R}_+ \times \Lambda \times \mathbb{R}_+^*.$$

Ainsi, le vecteur d'inconnues ϑ est de dimension $4 + 8 \times s$ et peut s'écrire sous la forme :

$$(\tau, p_b, p_r, p_v, \sigma_{1,1}^2, \rho_1, \eta_1, \sigma_{1,2}^2, \lambda_1, \mu_1(0), \mu_1(1), \sigma_{1,3}^2, \dots, \sigma_{s,1}^2, \rho_s, \eta_s, \sigma_{s,2}^2, \lambda_s, \mu_s(0), \mu_s(1), \sigma_{s,3}^2).$$

Nous donnons à présent l'expression de la matrice de transition a et des densités conditionnelles b^i (définies respectivement en (5.1) et (5.2)) dans le cadre précis de cet exemple.

Proposition 9. Un équivalent, lorsque $\Delta \rightarrow 0$, de la matrice $(a_{u,v}(\vartheta))_{u,v \in \mathbb{M}}$ est :

$$\begin{array}{c} \color{blue}{1} \qquad \qquad \qquad \color{red}{2} \qquad \qquad \qquad \color{green}{3} \\ \color{blue}{1} \left(\begin{array}{ccc} e^{-\tau\Delta} & p_b(1 - e^{-\tau\Delta}) & (1 - p_b)(1 - e^{-\tau\Delta}) \\ p_r(1 - e^{-\tau\Delta}) & e^{-\tau\Delta} & (1 - p_r)(1 - e^{-\tau\Delta}) \\ p_v(1 - e^{-\tau\Delta}) & (1 - p_v)(1 - e^{-\tau\Delta}) & e^{-\tau\Delta} \end{array} \right). \end{array}$$

Démonstration. On calcule un équivalent de $a_{u,v}(\vartheta)$ de la manière suivante. Chaque particule possède un taux de mutation constant égal à τ . Le temps d'attente entre deux mutations d'une particule suit alors une loi exponentielle $\mathcal{E}(\tau\Delta)$ et le nombre de mutations entre deux temps d'observation consécutifs suit par conséquent une loi de Poisson $\mathcal{P}(\tau\Delta)$.

Si $u = v$, on a par la formule des probabilités totales :

$$\begin{aligned} & \mathbb{P}(m_{t_k,i} = u | m_{t_{k-1},i} = u, z_{t_{k-1},i} = z) \\ &= \mathbb{P}(m_{t_k,i} = u | \{\text{pas de mutation entre } t_{k-1} \text{ et } t_k\}, m_{t_{k-1},i} = u, z_{t_{k-1},i} = z) \\ & \quad \times \mathbb{P}(\{\text{pas de mutation entre } t_{k-1} \text{ et } t_k\} | m_{t_{k-1},i} = u, z_{t_{k-1},i} = z) \\ & \quad + \mathbb{P}(m_{t_k,i} = u | \{\text{une mutation entre } t_{k-1} \text{ et } t_k\}, m_{t_{k-1},i} = u, z_{t_{k-1},i} = z) \\ & \quad \times \mathbb{P}(\{\text{une mutation entre } t_{k-1} \text{ et } t_k\} | m_{t_{k-1},i} = u, z_{t_{k-1},i} = z) \\ & \quad + \mathbb{P}(m_{t_k,i} = u, \{+ \text{ de 2 mutations entre } t_{k-1} \text{ et } t_k\} | m_{t_{k-1},i} = u, z_{t_{k-1},i} = z) \\ &= e^{-\tau\Delta} + 0 + \underset{\Delta \rightarrow 0}{o}(\Delta) \\ & \underset{\Delta \rightarrow 0}{\sim} e^{-\tau\Delta}. \end{aligned}$$

Si $u \neq v$, on a, en utilisant la formule des probabilités totales avec le même système complet d'événements que précédemment :

$$\begin{aligned} \mathbb{P}(m_{t_k,i} = v | m_{t_{k-1},i} = u, z_{t_{k-1},i} = z) &= 0 + \Delta\tau e^{-\tau\Delta} p_{u,v} + \underset{\Delta \rightarrow 0}{o}(\Delta) \\ & \underset{\Delta \rightarrow 0}{\sim} (1 - e^{-\tau\Delta}) p_{u,v}. \end{aligned}$$

□

La proposition suivante nous donne l'expression des fonctions b^i pour ce modèle. On rappelle que b^i est la densité de $z_{t_k,i}$ sachant $z_{t_{k-1},i} = y$ et $m_{t_k,i} = v$, pour tout

$i \in \llbracket 1, s \rrbracket$, pour tout $k \in \llbracket i_b + 1, i_d \rrbracket$ et pour tout $(y, v) \in \Lambda \times \mathbb{M}$.

Proposition 10. *Dans le modèle présenté ci-dessus, pour tout $i \in \llbracket 1, s \rrbracket$ la fonction b^i vaut :*

$$b^i(z|y, \ell; \vartheta) = \begin{cases} \phi(z, y, \sigma_{i,1}^2 \Delta) & \text{si } \ell = 1, \\ \phi\left(z, y + \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta, \sigma_{i,2}^2 \Delta\right) & \text{si } \ell = 2, \\ \phi\left(z, ye^{-\lambda_i \Delta} - \mu_i(1 - e^{-\lambda_i \Delta}), \sigma_{i,3}^2 v(\lambda_i)\right) & \text{si } \ell = 3, \end{cases}$$

où $v(\lambda_i) = \frac{1}{2\lambda_i} (1 - e^{-2\lambda_i \Delta})$ et $\phi(\cdot, \mu, \sigma^2)$ est la densité d'une loi normale $\mathcal{N}(\mu, \sigma^2 I_2)$ de dimension 2 de moyenne $\mu \in \mathbb{R}^2$ et de matrice de covariance $\sigma^2 I_2$, i.e. :

$$\phi(z, \mu, \sigma^2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|z - \mu\|^2}{2\sigma^2}\right).$$

Démonstration. Soit $i \in \llbracket 1, s \rrbracket$. On intègre les deux EDS qui dirigent les 2 types de mouvements (1 et 2) entre le temps t_k et le temps t_{k+1} avec $\Delta = t_{k+1} - t_k$:

- si $\ell = 1$: $X_{t_{k+1}} - X_{t_k} = \sigma_{i,1}(B_{t_{k+1}} - B_{t_k})$
- si $\ell = 2$: $X_{t_{k+1}} - X_{t_k} = \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta + \sigma_{i,2}(B_{t_{k+1}} - B_{t_k})$.

Or, comme $(B_{t_{k+1}} - B_{t_k}) \sim \mathcal{N}(0, \Delta I_2)$, on obtient que :

- si $\ell = 1$: $X_{t_{k+1}} - X_{t_k} \sim \mathcal{N}(0, \sigma_{i,1}^2 \Delta I_2)$
- si $\ell = 2$: $X_{t_{k+1}} - X_{t_k} \sim \mathcal{N}\left(\rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta, \sigma_{i,2}^2 \Delta I_2\right)$,

ce qui nous permet de déduire l'expression de b^i dans les cas où $\ell = 1$ ou $\ell = 2$.

Pour le mouvement de type 3, on se sert de la formule explicite de la solution du processus d'Ornstein-Uhlenbeck :

$$dX_t = -\lambda_i(X_t - \mu_i)dt + \sigma_{i,3}dB_t,$$

soit encore :

$$X_t = X_0 e^{-\lambda_i t} + \mu_i(1 - e^{-\lambda_i t}) + \int_0^t \sigma_{i,3} e^{\lambda_i(s-t)} dB_s.$$

Ainsi, on obtient

$$X_{t_{k+1}} = X_{t_k} e^{-\lambda_i \Delta} + \mu_i (1 - e^{-\lambda_i \Delta}) + \sigma_{i,3} e^{-\lambda_i t_{k+1}} \int_{t_k}^{t_{k+1}} e^{\lambda_i s} dB_s$$

et donc

$$X_{t_{k+1}} - X_{t_k} e^{-\lambda_i \Delta} - \mu_i (1 - e^{-\lambda_i \Delta}) = \sigma_{i,3} e^{-\lambda_i t_{k+1}} \int_{t_k}^{t_{k+1}} e^{\lambda_i s} dB_s.$$

On sait que ce terme de droite suit une loi normale centrée de variance :

$$\sigma_{i,3}^2 e^{-2\lambda_i t_{k+1}} \int_{t_k}^{t_{k+1}} e^{2\lambda_i s} ds = \frac{\sigma_{i,3}^2}{2\lambda_i} (1 - e^{-2\lambda_i \Delta}) = \sigma_{i,3}^2 v(\lambda_i).$$

Ainsi, on obtient l'expression de b^i annoncée dans la proposition. \square

5.3.2 Un argument du maximum de la fonction Q explicite dans cet exemple

Dans cette section, nous réécrivons la fonction $Q(\tilde{\vartheta}, \vartheta, \mathbf{Z})$ dans le cas particulier de l'exemple traité dans cette section, afin de trouver l'argument de son maximum.

On montre dans la section F.4 de l'annexe F, que les conditions **(C1)**, **(C2)** et **(C3)** sont satisfaites, et la proposition 6 nous donne ainsi la forme de la fonction Q .

Par cette proposition 6, et comme a ne dépend pas de z , on a :

$$\begin{aligned} Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \sum_{u=1}^M \sum_{v=1}^M \xi_{k-1}^i(u, v; \vartheta) \ln(a_{u,v}(\tilde{\vartheta})) \\ &\quad + \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \sum_{v=1}^M \gamma_k^i(v; \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, v, \tilde{\vartheta} \right) \right) + C(\vartheta, \mathbf{Z}, \Delta) + \underset{\Delta \rightarrow 0}{o}(1). \end{aligned}$$

Afin de mettre en place l'étape M de l'algorithme EM, on recherche l'argument du maximum selon la variable $\tilde{\vartheta}$ de $Q(\tilde{\vartheta}, \vartheta, \mathbf{Z})$.

Il est établi que :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = Q^1(\tilde{\vartheta}, \vartheta, \mathbf{Z}) + Q^2(\tilde{\vartheta}, \vartheta, \mathbf{Z}) + C(\vartheta, \mathbf{Z}, \Delta) + \underset{\Delta \rightarrow 0}{o}(1),$$

où :

$$Q^1(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \sum_{u=1}^M \sum_{v=1}^M \xi_{k-1}^i(u, v; \vartheta) \ln(a_{u,v}(\tilde{\vartheta})),$$

$$Q^2(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \sum_{v=1}^M \gamma_k^i(v; \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, v, \tilde{\vartheta} \right) \right).$$

Afin de calculer $\arg \max_{\tilde{\vartheta}} Q(\tilde{\vartheta}, \vartheta, \mathbf{Z})$, nous faisons remarquer que les différents paramètres inconnus de $\tilde{\vartheta}$ ne sont pas localisés aux mêmes endroits, ainsi la recherche de l'argument du maximum de $Q(\cdot, \vartheta, \mathbf{Z})$ selon les variables (τ, p_b, p_r, p_v) , qui n'interviennent dans Q que via a revient à la recherche de l'argument du maximum de $Q^1(\cdot, \vartheta, \mathbf{Z})$. De même, la recherche de l'argument du maximum de $Q(\cdot, \vartheta, \mathbf{Z})$ selon les variables $(\sigma_{i,1}^2, \rho_i, \eta_i, \sigma_{i,2}^2, \lambda_i, \mu_i(0), \mu_i(1), \sigma_{i,3}^2)$, qui n'interviennent dans Q que via b^i revient à rechercher l'argument du maximum de $Q^2(\cdot, \vartheta, \mathbf{Z})$.

5.3.2.1 Recherche de l'argument du maximum selon les paramètres τ, p_b, p_r, p_v

Proposition 11. *L'argument du maximum de $Q(\cdot, \vartheta, \mathbf{Z})$ selon les variables*

$$(\tau, p_b, p_r, p_v) \in \mathbb{R}_+ \times [0, 1] \times [0, 1] \times [0, 1]$$

est atteint pour :

$$\tau = \frac{1}{\Delta} \ln \left(1 + \frac{\sum_{i=1}^s \sum_{k=2}^{n_i} \sum_{m \neq n} \xi_{k-1}^i(m, n; \vartheta)}{\sum_{i=1}^s \sum_{k=2}^{n_i} \sum_{m=1}^3 \xi_{k-1}^i(m, m; \vartheta)} \right),$$

$$p_b = \frac{\sum_{i=1}^s \sum_{k=2}^{n_i} \xi_{k-1}^i(1, 2; \vartheta)}{\sum_{i=1}^s \sum_{k=2}^{n_i} \xi_{k-1}^i(1, 2; \vartheta) + \sum_{i=1}^s \sum_{k=2}^{n_i} \xi_{k-1}^i(1, 3; \vartheta)},$$

$$p_r = \frac{\sum_{i=1}^s \sum_{k=2}^{n_i} \xi_{k-1}^i(2, 1; \vartheta)}{\sum_{i=1}^s \sum_{k=2}^{n_i} \xi_{k-1}^i(2, 1; \vartheta) + \sum_{i=1}^s \sum_{k=2}^{n_i} \xi_{k-1}^i(2, 3; \vartheta)},$$

$$p_v = \frac{\sum_{i=1}^s \sum_{k=2}^{n_i} \xi_{k-1}^i(3, 1; \vartheta)}{\sum_{i=1}^s \sum_{k=2}^{n_i} \xi_{k-1}^i(3, 1; \vartheta) + \sum_{i=1}^s \sum_{k=2}^{n_i} \xi_{k-1}^i(3, 2; \vartheta)}.$$

Démonstration. La preuve de cette proposition est reportée dans la section G.1 de l'annexe G. \square

5.3.2.2 Si les paramètres qui interviennent dans les mouvements sont propres à chaque trajectoire

Dans cette sous-partie, on suppose que les paramètres $\sigma_{i,1}^2$, ρ_i , η_i , $\sigma_{i,2}^2$, λ_i , $\mu_i(0)$, $\mu_i(1)$, $\sigma_{i,3}^2$ sont individualisés pour chaque trajectoire, c'est-à-dire que chaque trajectoire \mathbf{Z}^i a 8 paramètres qui lui sont propres et supposés inconnus.

On recherche l'argument du maximum selon les variables $(\sigma_{i,1}^2, \rho_i, \eta_i, \sigma_{i,2}^2, \lambda_i, \mu_i(0), \mu_i(1), \sigma_{i,3}^2)$ de :

$$Q((\tau, p_b, p_r, p_v, (\sigma_{1,1}^2, \rho_1, \eta_1, \sigma_{1,2}^2, \lambda_1, \mu_1(0), \mu_1(1), \sigma_{1,3}^2), \dots, (\sigma_{s,1}^2, \rho_s, \eta_s, \sigma_{s,2}^2, \lambda_s, \mu_s(0), \mu_s(1), \sigma_{s,3}^2)), \vartheta, \mathbf{Z}).$$

Proposition 12. *L'argument du maximum de $Q(\cdot, \vartheta, \mathbf{Z})$ selon les variables*

$$(\sigma_{i,1}^2, \rho_i, \eta_i, \sigma_{i,2}^2, \lambda_i, \mu_i, \sigma_{i,3}^2) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times [0, 2\pi[\times \mathbb{R}_+^* \times \mathbb{R}_+ \times \Lambda \times \mathbb{R}_+^*$$

est atteint pour :

$$\sigma_{i,1}^2 = \frac{1}{2\Delta} \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(1; \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(1; \vartheta)},$$

$$\eta_i = \arctan \left(\frac{\sum_{k=i_b+1}^{i_d} (z_{t_k,i}(1) - z_{t_{k-1},i}(1)) \gamma_k^i(2; \vartheta)}{\sum_{k=i_b+1}^{i_d} (z_{t_k,i}(0) - z_{t_{k-1},i}(0)) \gamma_k^i(2; \vartheta)} \right),$$

$$\rho_i = \frac{1}{\Delta} \frac{\sum_{k=i_b+1}^{i_d} \left\langle \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix}, z_{t_k,i} - z_{t_{k-1},i} \right\rangle \gamma_k^i(2; \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(2; \vartheta)},$$

$$\sigma_{i,2}^2 = \frac{1}{2\Delta} \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(2; \vartheta) \left\| z_{t_k,i} - z_{t_{k-1},i} - \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta \right\|^2}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(2; \vartheta)},$$

$$\mu_i = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) (z_{t_k,i} - z_{t_{k-1},i} a_i)}{(1 - a_i) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta)},$$

$$\sigma_{i,3}^2 = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) \|z_{t_k,i} - z_{t_{k-1},i} a_i - \mu_i (1 - a_i)\|^2}{2v(\lambda_i) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta)},$$

$$\lambda_i = -\frac{1}{\Delta} \ln(a_i) \quad \text{où} \quad a_i = \begin{cases} \hat{a}_i & \text{si } \hat{a}_i > 0 \\ 0 & \text{si } \hat{a}_i \leq 0 \end{cases} \quad \text{avec}$$

$$\hat{a}_i = \frac{\left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) \right) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) \langle z_{t_{k-1},i}, z_{t_k,i} \rangle - \left\langle \sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) z_{t_{k-1},i}, \sum_{j=i_b+1}^{i_d} \gamma_j^i(3; \vartheta) z_{t_j,i} \right\rangle}{\left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) \right) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) \|z_{t_{k-1},i}\|^2 - \left\| \sum_{j=i_b+1}^{i_d} \gamma_j^i(3; \vartheta) z_{t_{j-1},i} \right\|^2}.$$

Démonstration. La preuve de cette proposition est reportée dans la section G.2 de l'annexe G. \square

Remarque : Si $a_i = 0$, alors $\lambda_i = +\infty$, et le maximum n'est pas atteint ; il s'agit en fait d'une borne supérieure.

5.3.2.3 Si les paramètres $\sigma_{i,1}$, $\sigma_{i,2}$, $\sigma_{i,3}$, ρ_i et λ_i sont communs à toutes les trajectoires

Dans cette sous-partie, on suppose que certains paramètres des trois mouvements sont communs à toutes les trajectoires. On va supposer que pour tout $i \in \llbracket 1, s \rrbracket$, $\sigma_{i,1}^2 = \sigma_b^2$, $\rho_i = \rho$, $\sigma_{i,2}^2 = \sigma_r^2$, $\lambda_i = \lambda$, et $\sigma_{i,3}^2 = \sigma_v^2$, c'est-à-dire que tous les morceaux de trajectoires de type 1 ont le même coefficient de diffusion σ_b , ceux de type 2 ont le même coefficient de diffusion σ_r et le même drift v et ceux de type 3 ont le même coefficient de diffusion σ_v et la même force de rappel λ .

Ainsi, le paramètre inconnu ϑ n'est plus de dimension $4 + 8 \times s$ comme précédemment mais de dimension $9 + 3p$ et on l'écrit sous la forme :

$$(\tau, p_b, p_r, p_v, \sigma_b^2, \rho, \sigma_r^2, \lambda, \sigma_v^2, \eta_1, \mu_1, \dots, \eta_s, \mu_s) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+ \times \mathbb{R}_+^* \times ([0, 2\pi[\times\Lambda)^s.$$

Comme précédemment, on cherche $\arg \max_{\tilde{\vartheta}} Q(\tilde{\vartheta}, \vartheta, \mathbf{Z})$, en cherchant séparément, grâce à la forme de Q , l'argument du maximum selon les variables $(\tau, p_b, p_r, p_v) \in \mathbb{R}_+^4$, puis celui selon les variables $(\sigma_b^2, \rho, \sigma_r^2, \lambda, \sigma_v^2, \eta_1, \mu_1, \dots, \eta_s, \mu_s) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+ \times \mathbb{R}_+^* \times ([0, 2\pi[\times\Lambda)^s$.

La recherche de l'argument du maximum selon les variables (τ, p_b, p_r, p_v) se fait exactement de la même manière que ce qui a été fait dans la proposition 11. Cherchons maintenant celui selon les variables $(\sigma_b^2, \rho, \sigma_r^2, \lambda, \sigma_v^2, \eta_1, \mu_1, \dots, \eta_s, \mu_s)$.

Proposition 13. *L'argument du maximum de $Q(\cdot, \vartheta, \mathbf{Z})$ selon les variables*

$$(\sigma_b^2, \rho, \sigma_r^2, \lambda, \sigma_v^2, \eta_1, \mu_1, \dots, \eta_s, \mu_s) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+ \times \mathbb{R}_+^* \times ([0, 2\pi[\times\Lambda)^s$$

est atteint pour les mêmes valeurs que dans la proposition 12 pour η_i et μ_i . Pour les

autres paramètres, elle est atteinte en :

$$\sigma_b^2 = \frac{1}{2\Delta} \frac{\sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(1; \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2}{\sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(1; \vartheta)},$$

$$\rho = \frac{1}{\Delta} \frac{\sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \left\langle \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix}, z_{t_k, i} - z_{t_{k-1}, i} \right\rangle \gamma_k^i(2; \vartheta)}{\sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(2; \vartheta)},$$

$$\sigma_r^2 = \frac{1}{2\Delta} \frac{\sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(2; \vartheta) \left\| z_{t_k, i} - z_{t_{k-1}, i} - \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta \right\|^2}{\sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(2; \vartheta)},$$

$$\sigma_v^2 = \frac{\sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} a_i - \mu_i(1-a)\|^2}{2v(\lambda) \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta)},$$

$$\lambda = -\frac{1}{\Delta} \ln(a) \quad \text{où} \quad a = \begin{cases} \hat{a} & \text{si } \hat{a} > 0 \\ 0 & \text{si } \hat{a} \leq 0 \end{cases} \quad \text{avec}$$

$$\hat{a} = \frac{\sum_{i=1}^s \left[\sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) \langle z_{t_{k-1}, i}, z_{t_k, i} \rangle - \left\langle \sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) z_{t_{k-1}, i}, \frac{\sum_{j=i_b+1}^{i_d} \gamma_j^i(3; \vartheta) z_{t_j, i}}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(3; \vartheta)} \right\rangle \right]}{\sum_{i=1}^s \left[\sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) \|z_{t_{k-1}, i}\|^2 - \frac{\left\| \sum_{j=i_b+1}^{i_d} \gamma_j^i(3; \vartheta) z_{t_{j-1}, i} \right\|^2}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(3; \vartheta)} \right]}.$$

Démonstration. La preuve de cette proposition est reportée dans la section G.3 de

l'annexe G. □

Remarque : Comme dans la proposition précédente, si $a = 0$, alors $\lambda = +\infty$, et le maximum n'est pas atteint ; il s'agit en fait d'une borne supérieure.

Si on suppose de plus que les trois mouvements considérés ont les mêmes coefficients de diffusion ($\sigma_b = \sigma_r = \sigma_v = \sigma$), alors on obtient la proposition suivante.

Proposition 14. *Si $\sigma_b = \sigma_r = \sigma_v = \sigma$, alors l'argument du maximum de $Q(\cdot, \vartheta, \mathbf{Z})$ selon les variables*

$$(\sigma^2, \rho, \lambda, \eta_1, \mu_1, \dots, \eta_s, \mu_s) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+ \times ([0, 2\pi[\times\Lambda)^s$$

est atteint pour les mêmes valeurs que dans la proposition 13 pour ρ, λ, η_i et μ_i . Pour σ , il est atteint en :

$$\sigma^2 = \frac{n_\sigma}{\sum_{j=1}^s \sum_{k=j_b+1}^{j_d} (\gamma_k^j(1; \vartheta) + \gamma_k^j(2; \vartheta) + \gamma_k^j(3; \vartheta))}$$

avec

$$\begin{aligned} n_\sigma &= \frac{1}{2\Delta} \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(1; \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2 \\ &+ \frac{1}{2\Delta} \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(2; \vartheta) \left\| z_{t_k, i} - z_{t_{k-1}, i} - v \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta \right\|^2 \\ &+ \frac{1}{2v(\lambda)} \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(3; \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} a - \mu_i(1-a)\|^2. \end{aligned}$$

Démonstration. La preuve de cette proposition est reportée dans la section G.4 de l'annexe G. □

5.4 Conclusion

Ce chapitre avait pour but d'élaborer une méthode permettant d'estimer les classes de mouvement pour plusieurs trajectoires dans un domaine donné, en tenant compte de la possibilité de changer de mouvement lors du déplacement des particules. La méthode que nous avons choisie se base sur un modèle de Markov à états cachés et sur l'algorithme EM pour calculer l'estimateur au sens du maximum de vraisemblance des caractéristiques des mouvements. Dans ce chapitre, nous avons établi les formules explicites des solutions au sens du maximum de vraisemblance dans le cas où les trois mouvements envisagés sont respectivement le mouvement brownien, le mouvement brownien dirigé et le processus d'Ornstein-Uhlenbeck. Dans le chapitre suivant, nous présentons la validation de notre méthode sur des données simulées et des données réelles.

APPLICATIONS SUR DES DONNÉES SIMULÉES PUIS SUR DES DONNÉES RÉELLES

Dans ce chapitre, nous analysons un jeu de trajectoires simulées avec la théorie présentée dans le chapitre précédent dans le but de déterminer les paramètres inconnus qui régissent le mouvement de ces trajectoires. Dans un deuxième temps, nous nous servirons des résultats obtenus sur les données simulées pour valider notre méthode sur un jeu de trajectoires issues de données réelles. Les données réelles considérées seront les mêmes qu'aux chapitres 2 et 4.

6.1 Résultats sur des données simulées

Nous recourons à l'algorithme EM en considérant certains paramètres tantôt comme connus, tantôt comme inconnus. Nous utilisons chaque fois le même jeu de trajectoires, présenté ci-dessous.

Le jeu est composé de $s = 4000$ trajectoires qui proviennent du résultat de la simulation d'un processus BDMM tel que celui présenté dans la section 5.3.1 du chapitre 5, et résumé ici :

- La marque de chaque particule identifie son régime de mouvement à chaque instant, qui peut être de trois types différents (décrits ci-dessous). L'ensemble \mathbb{M} des marques est :

$$\mathbb{M} = \{1, 2, 3\}.$$

- Chaque particule possède un taux de mutation constant égal à τ :

$$\tau(\mathbf{x}) = \tau \times n(\mathbf{x}).$$

- Le noyau de mutation est celui de l'exemple 3 du chapitre 3. Pour $(\mathbf{x}, \mathbf{y}) \in E^2$, on a

$$k_\tau(\mathbf{x}, \mathbf{y}) = \begin{cases} \frac{1}{n(\mathbf{x})} p_{m_i, m} & \text{si } \exists x_i = (z_i, m_i) \in \mathbf{x}, \exists m \in \mathbb{M}, \mathbf{y} = (\mathbf{x} \setminus (z_i, m_i)) \cup (z_i, m), \\ 0 & \text{sinon,} \end{cases}$$

où $(p_{i,j})_{(i,j) \in \{1,2,3\}^2}$ est une matrice de transition de la forme

$$\begin{pmatrix} 0 & p_b & 1 - p_b \\ p_r & 0 & 1 - p_r \\ p_v & 1 - p_v & 0 \end{pmatrix}.$$

- Les particules changent de position chacune selon sa marque et indépendamment des autres.
- Le noyau de naissance est une loi uniforme sur le carré $\Lambda := [0, 250] \times [0, 250]$. À chaque naissance, la couleur est choisie de manière équiprobable entre les trois couleurs 1, 2, 3, i.e. :

$$k_\beta(\mathbf{x}, \mathbf{y}) = \begin{cases} \frac{1}{3} \times \frac{1}{250^2} \mathbb{1}_\Lambda(z) & \text{si } \exists (z, m) \in \Lambda \times \mathbb{M}, \mathbf{y} = \mathbf{x} \cup (z, m) \\ 0 & \text{sinon.} \end{cases}$$

- Le noyau de mort est uniforme sur l'ensemble des particules présentes.
- Les taux de naissance et de mort sont constants, égaux à 300 et β vérifie :

$$\exists n^* \in \mathbb{N}, \forall \mathbf{x} \in E, n(\mathbf{x}) \geq n^* \implies \beta(\mathbf{x}) = 0.$$

Les trois types de mouvement 1, 2 ou 3 sont ceux détaillés dans la section 5.3 du chapitre 5, à savoir pour chaque trajectoire $i \in \llbracket 1, s \rrbracket$:

- **Type 1** : mouvement brownien de coefficient de diffusion $\sigma_{i,1}$.
- **Type 2** : mouvement brownien de coefficient de diffusion $\sigma_{i,2}$ avec drift égal à $\rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix}$.
- **Type 3** : mouvement qui suit un processus d'Ornstein-Uhlenbeck de coefficient de diffusion $\sigma_{i,3}$, de force de rappel λ_i et de point d'équilibre μ_i .

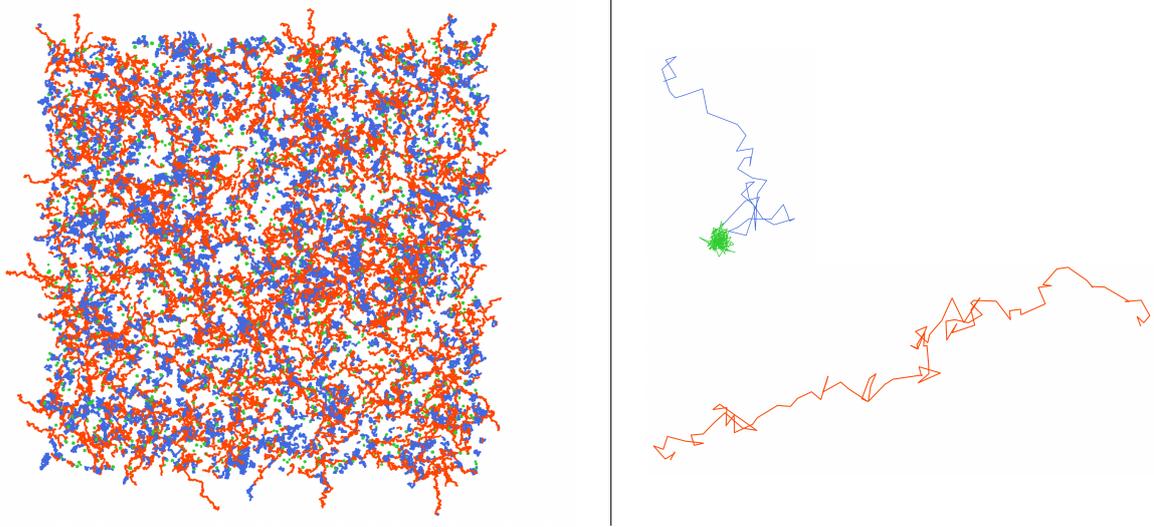


FIGURE 6.1 – A gauche : 4000 trajectoires comportant à elles toutes 1071 changements de types de mouvement. A droite : Deux trajectoires (zoomées) parmi les 4000 de gauche : la 138-ième (en haut) de longueur 177 points qui suit le type de mouvement 1 pendant 43 pas de temps puis qui change de type de mouvement pour suivre le type 3, la 301-ième (en bas) est une trajectoire de longueur 98 points qui suit le type de mouvement 2 durant toute sa durée de vie.

Pour la simulation, les paramètres sont les suivants : $\sigma_{i,1} = 1.22$, $\rho_i = 1.44$, $\sigma_{i,2} = 1.2$, $\lambda_i = 6$, $\sigma_{i,3} = 0.8$ pour tout $i \in \llbracket 1, s \rrbracket$. Quant à l'angle η_i du type de mouvement 2, il est tiré selon une loi uniforme entre 0 et 2π pour chaque trajectoire i . Pour le point d'équilibre $\mu_i = \begin{pmatrix} \mu_i(0) \\ \mu_i(1) \end{pmatrix}$ lié au type 3, on prend la position dans laquelle se trouve la particule i au moment où elle initie un déplacement dans le type 3. L'inter-temps Δ vaut 0.14. On fixe aussi les paramètres avec les valeurs suivantes : $\tau = 0.04$, $p_b = 0.25$, $p_v = 0.5$, $p_r = 1$.

Avec tous ces paramètres, la proposition 9 nous donne que la matrice $\{a_{u,v}(\vartheta)\}_{(u,v) \in \mathbb{M}^2}$ vaut, en arrondissant à 10^{-3} près :

$$a = \begin{pmatrix} e^{-\tau\Delta} & p_b(1 - e^{-\tau\Delta}) & (1 - p_b)(1 - e^{-\tau\Delta}) \\ p_r(1 - e^{-\tau\Delta}) & e^{-\tau\Delta} & (1 - p_r)(1 - e^{-\tau\Delta}) \\ p_v(1 - e^{-\tau\Delta}) & (1 - p_v)(1 - e^{-\tau\Delta}) & e^{-\tau\Delta} \end{pmatrix} = \begin{pmatrix} 0.994 & 0.004 & 0.002 \\ 0.006 & 0.994 & 0 \\ 0.003 & 0.003 & 0.994 \end{pmatrix}.$$

La Figure 6.1 affiche les trajectoires issues du résultat d'une telle simulation avec en bleu les morceaux de trajectoires qui suivent le type de mouvement 1, en rouge ceux qui suivent le type 2, et en vert ceux qui suivent le type 3.

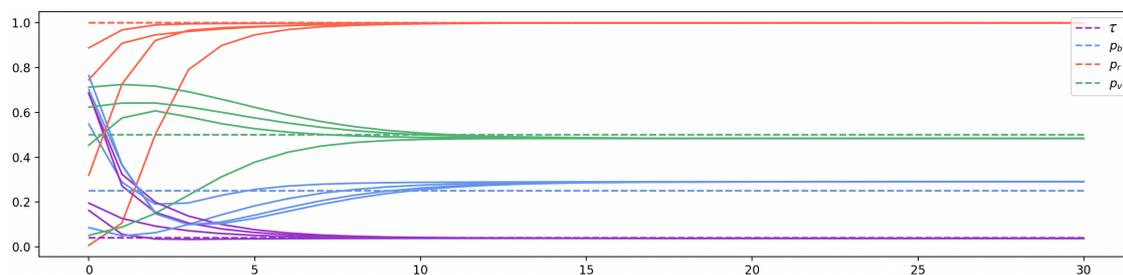


FIGURE 6.2 – Évolution du maximum de vraisemblance estimé par l'algorithme EM à chaque itération en fonction du nombre d'itérations. On suppose que seuls les paramètres τ , p_b , p_r et p_v sont inconnus. L'algorithme est exécuté pour 4 quadruplets de valeurs initiales différents, choisis selon une loi uniforme sur $[0, 1]^4$.

6.1.1 Cas où seuls les paramètres de la matrice de transition a sont inconnus

Dans cette section, on teste l'algorithme EM présenté en section 5.2.2.1 en considérant que les paramètres $\sigma_{i,1}$, ρ_i , $\sigma_{i,2}$, λ_i , $\sigma_{i,3}$, μ_i et η_i sont connus. Ils ne feront donc pas partie de la variable inconnue ϑ . Les paramètres inconnus sont les paramètres de la matrice a , c'est-à-dire τ , p_b , p_r et p_v .

Comme expliqué ci-dessus, pour pouvoir utiliser l'algorithme EM, il faut initialiser les valeurs des paramètres recherchés. On sait ici que $0 \leq p_b, p_r, p_v \leq 1$ car sont des probabilités et que τ est un réel positif qui a plutôt tendance à être proche de 0 (sinon les changements de type de mouvement auraient lieu extrêmement souvent). On va donc fournir comme valeur initiale à l'algorithme EM pour ces 4 paramètres, 4 réels tirés selon une loi uniforme sur $[0, 1]$. De plus, comme l'algorithme EM converge vers un minimum local, nous allons l'exécuter plusieurs fois, avec plusieurs valeurs initiales choisies aléatoirement dans $[0, 1]$.

Le graphique de la figure 6.2 reporte pour 4 quadruplets de valeurs initiales différents le maximum renvoyé par l'algorithme EM à chaque itération, en fonction du nombre d'itérations de l'algorithme. On arrête l'algorithme au bout de 30 itérations. On remarque que dans les 4 cas, l'algorithme EM converge vers la même valeur. Le choix du quadruplet initial n'influence donc pas sur le résultat, ce qui est cohérent

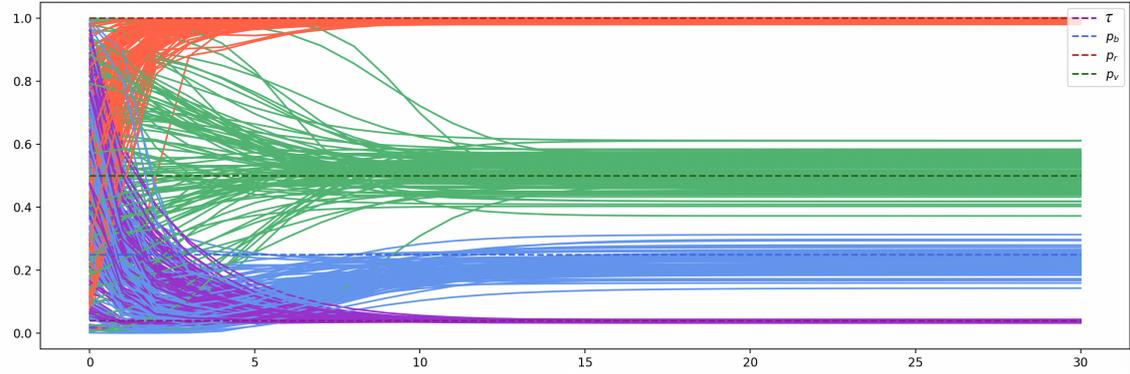


FIGURE 6.3 – Évolution du maximum de vraisemblance estimé par l'algorithme EM à chaque itération en fonction du nombre d'itérations pour 100 expériences similaires à celle utilisée pour la figure 6.2.

avec la non existence de maxima locaux (cf preuve de la proposition 11 du chapitre 5).

Dans un deuxième temps, l'algorithme EM est exécuté sur 100 expériences réalisées dans les mêmes conditions. Étant donné que les différents quadruplets initiaux mènent au même résultat, j'ai choisi pour chaque expérience, un seul quadruplet initial aléatoire dans $[0, 1]^4$. La figure 6.3 regroupe le tracé de l'Évolution du maximum de vraisemblance estimé par l'algorithme EM à chaque itération en fonction du nombre d'itérations pour les 100 expériences. Cette figure montre la convergence de l'algorithme EM pour les 100 expériences et nous conforte quant à l'efficacité de l'algorithme, attesté par la faible variance du résultat.

6.1.2 Cas où les paramètres de chaque trajectoire sont inconnus et distincts

Dans cette section, on suppose connus les paramètres τ , p_b , p_r et p_v de la matrice de transition a qui régit les changements de type de mouvement. Estimer les paramètres $\sigma_{i,1}$, ρ_i , $\sigma_{i,2}$, λ_i , $\sigma_{i,3}$, μ_i et η_i revient donc à s'intéresser à chaque trajectoire de manière indépendante. On va donc se concentrer sur une seule trajectoire i par la suite.

Pour la simulation, on conserve les valeurs précédentes pour les paramètres :

$\sigma_{i,1} = 1.22$, $\rho_i = 1.44$, $\sigma_{i,2} = 1.2$, $\lambda_i = 6$, $\sigma_{i,3} = 0.8$ et $\tau = 0.04$, $p_b = 0.25$, $p_v = 0.5$, $p_r = 1$. Comme précédemment, pour la simulation, on choisit pour l'angle η_i un nombre tiré uniformément entre 0 et 2π et pour le point d'équilibre μ_i , la position dans laquelle se trouve la particule i au moment où elle initie un déplacement dans le type de mouvement **3** pour μ_i .

Pour tester l'algorithme EM, on a besoin de fixer des valeurs initiales pour les paramètres inconnus. On choisit donc :

- pour $\sigma_{i,1}$, $\sigma_{i,2}$ et $\sigma_{i,3}$ un nombre tiré selon une loi uniforme sur $[0, 3]$,
- pour ρ_i un nombre tiré selon une loi uniforme sur $[0, 10]$,
- pour λ_i un nombre tiré selon une loi uniforme sur $[0, 10]$,
- pour μ_i un point qui est sur la trajectoire i ,
- pour η_i l'angle décrit par le vecteur qui relie le premier et le dernier point de la trajectoire.

Nous testons l'algorithme EM sur les deux trajectoires particulières présentées à droite de la Figure 6.1 en prenant pour valeurs initiales de l'algorithme pour le paramètre μ_i , soit le premier point, soit le point médian, soit le dernier point de la trajectoire, et pour les paramètres $(\sigma_{i,1}, \rho_i, \sigma_{i,2}, \lambda_i, \sigma_{i,3})$, soit θ_1 , soit θ_2 , soit θ_3 , soit θ_4 avec :

$$\begin{aligned}\theta_1 &= (1.86, 7.08, 0.82, 3.75, 1.77), \\ \theta_2 &= (0.45, 4.54, 1.27, 2.56, 2.88), \\ \theta_3 &= (1.89, 8.40, 1.33, 4.07, 0.70), \\ \theta_4 &= (1, 2, 1, 2, 1).\end{aligned}$$

Au total, cela fait donc $3 \times 4 = 12$ valeurs initiales différentes. Le tableau 6.1 fait correspondre le numéro de l'exécution avec les valeurs initiales choisies.

Avec ces 12 points de départ, il est difficile d'afficher, comme dans la section précédente, les graphes de convergence en fonction des itérations de l'algorithme. En effet, suivant la valeur initiale choisie, les paramètres ne convergent pas vers les mêmes valeurs, donc ce type de graphe devient très rapidement illisible. On choisira plutôt d'estimer, pour chaque valeur initiale, les états cachés (i.e. les types de mouvement **1**, **2** ou **3**) de chaque point de la trajectoire étudiée grâce à l'algorithme

Numéro de l'exécution de l'algo EM	Valeur initiale pour le paramètre μ_i	Valeur initiale pour les paramètres $(\sigma_{i,1}, \rho_i, \sigma_{i,2}, \lambda_i, \sigma_{i,3})$
1	Premier point de la trajectoire	θ_4
2	Point médian de la trajectoire	
3	Dernier point de la trajectoire	
4	Premier point de la trajectoire	θ_1
5		θ_2
6		θ_3
7	Point médian de la trajectoire	θ_1
8		θ_2
9		θ_3
10	Dernier point de la trajectoire	θ_1
11		θ_2
12		θ_3

TABLE 6.1 – Valeurs initiales des 12 exécutions de l'algorithme EM.

de Viterbi (proposition 8, en prenant pour valeur des paramètres celle obtenue à la dernière itération de l'algorithme EM) puis d'afficher la trajectoire, coloriée en fonction de ces états cachés estimés. Nous affichons aussi pour chaque point de départ la valeur de la vraisemblance calculée en ce point. Les figures 6.4 et 6.5 regroupent ces différents résultats pour les deux trajectoires de la figure 6.1.

Remarques sur les figures 6.4 et 6.5 : On observe, en comparant les résultats obtenus avec les valeurs initiales 7, 8, 9, 10, 11 et 12 dans la figure 6.5, que les valeurs initiales de $\sigma_{i,1}$, ρ_i , $\sigma_{i,2}$, λ_i , $\sigma_{i,3}$ ont peu de répercussion sur le résultat. En revanche la valeur initiale donnée au point d'équilibre μ_i a quant à elle beaucoup d'influence sur le résultat (exécutions 1, 2, 3).

Au contraire, d'après la figure 6.4, la valeur initiale du point d'équilibre μ_i n'a pas beaucoup d'incidence sur les résultats (exécutions 1, 2, 3) alors que les valeurs initiales de $\sigma_{i,1}$, ρ_i , $\sigma_{i,2}$, λ_i , $\sigma_{i,3}$ ont quant à elles beaucoup d'influence sur le résultat (exécutions 4 à 12).

On choisit comme "meilleure" valeur initiale celle qui correspond à la plus haute

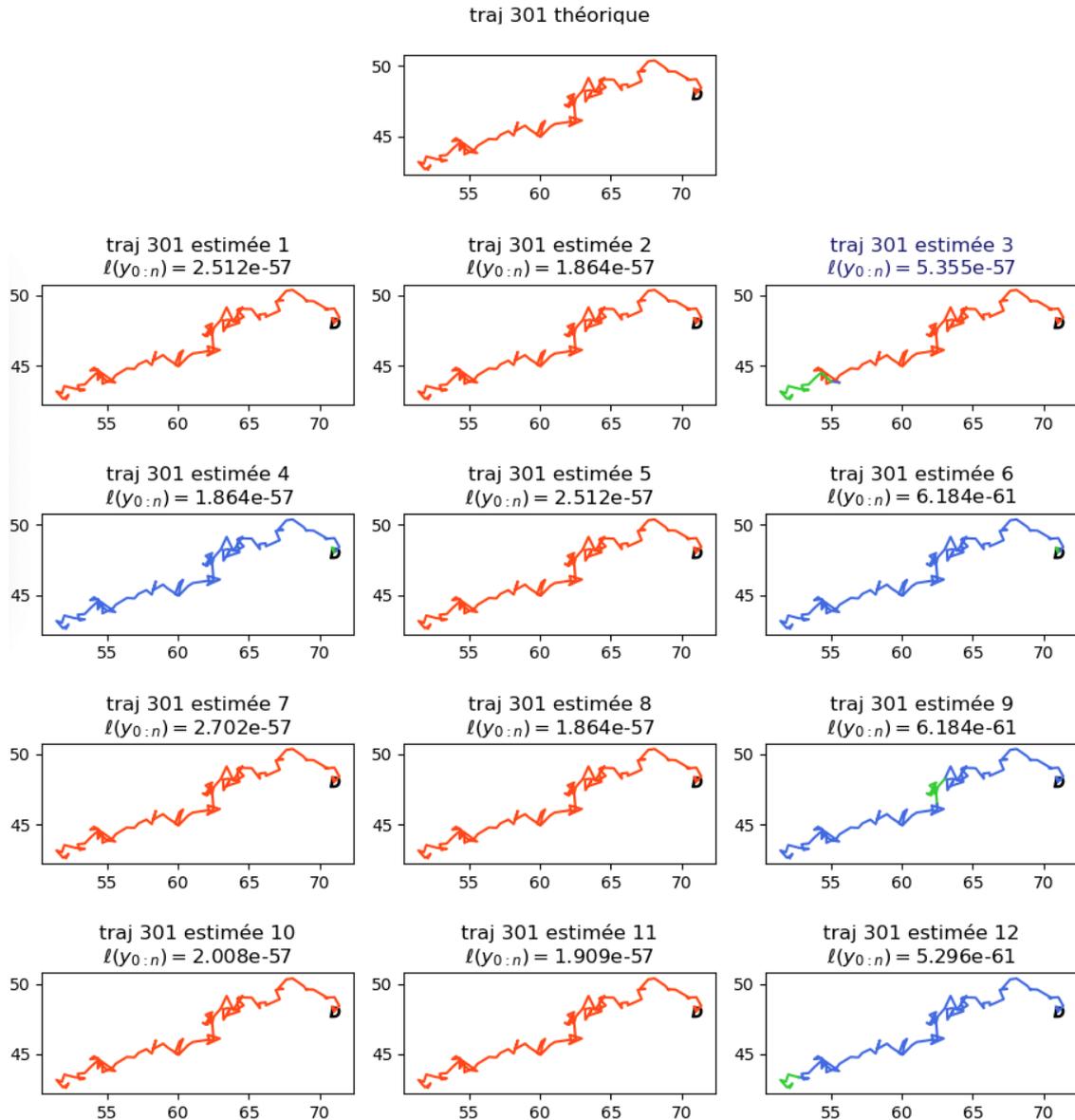


FIGURE 6.4 – Résultat de l'estimation des états cachés pour la trajectoire 301 de l'expérience de la figure 6.1, basée sur les paramètres estimés grâce à l'algorithme EM avec 12 points de départ différents. La valeur au dessus de chaque graphique est la valeur de la log-vraisemblance obtenue avec les paramètres estimés. Le point de départ de la trajectoire est noté **D**.

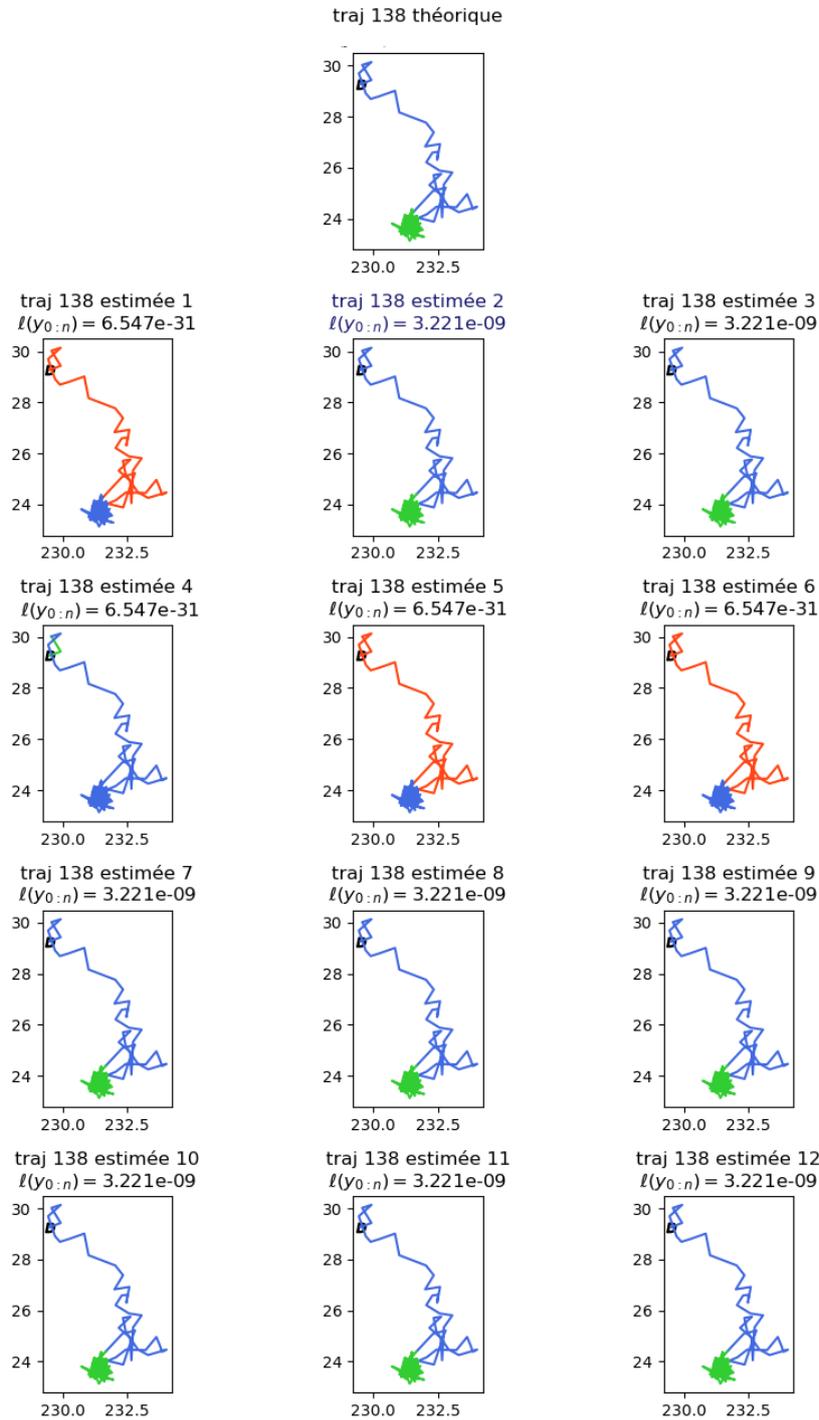


FIGURE 6.5 – Résultat de l'estimation des états cachés pour la trajectoire 138 de l'expérience de la figure 6.1, basée sur les paramètres estimés grâce à l'algorithme EM avec 12 points de départ différents. La valeur au dessus de chaque graphique est la valeur de la log-vraisemblance obtenue avec les paramètres estimés. Le point de départ de la trajectoire est noté **D**.

valeur de vraisemblance, c'est-à-dire ici la valeur initiale 3 pour la trajectoire 301 et au point de départ 2 pour la trajectoire 138.

On ne va pas davantage s'attarder sur cette étude car si les paramètres sont différents pour chaque trajectoire, alors nous sommes confrontés à un problème d'identifiabilité des paramètres. En effet, il n'est pas aisé de distinguer une trajectoire qui suit le type 1 d'une trajectoire qui suit le type 2 avec un drift très faible. De la même manière, une trajectoire qui suit un mouvement de type 3 avec un coefficient de diffusion élevé et une faible force de rappel vers le point d'équilibre est similaire à une trajectoire qui suit le type de mouvement 1 avec un coefficient de diffusion plus faible.

On constate donc que si les paramètres sont différents pour chaque trajectoire alors l'algorithme EM a déjà beaucoup de mal à les estimer, même en supposant connus les autres paramètres. Ainsi, si on espère pouvoir mettre en place un algorithme EM efficace pour estimer tous les paramètres en même temps, on ne peut pas supposer les paramètres $\sigma_{i,1}$, ρ_i , $\sigma_{i,2}$, λ_i , $\sigma_{i,3}$ différents pour chaque trajectoire. Par conséquent, dans la section suivante, nous traitons le cas où les paramètres $\sigma_{i,1}$, ρ_i , $\sigma_{i,2}$, λ_i , $\sigma_{i,3}$ sont supposés identiques à toutes les trajectoires \mathbf{Z}^i , $i \in \llbracket 1, s \rrbracket$.

6.1.3 Cas où les paramètres $\sigma_{i,1}$, ρ_i , $\sigma_{i,2}$, λ_i , $\sigma_{i,3}$ sont identiques pour toutes les trajectoires

On suppose ici, comme dans la partie 5.3.2.3, que les coefficients de diffusions de trois types de mouvement, la force de rappel du types 3 et la norme du vecteur drift du type 2, c'est-à-dire les paramètres $\sigma_{i,1}$, ρ_i , $\sigma_{i,2}$, λ_i , $\sigma_{i,3}$ sont identiques à toutes les trajectoires. Pour tout $i \in \llbracket 1, s \rrbracket$, on a :

$$\left\{ \begin{array}{l} \sigma_{i,1} = \sigma_b, \\ \rho_i = \rho, \\ \sigma_{i,2} = \sigma_r, \\ \lambda_i = \lambda, \\ \sigma_{i,3} = \sigma_v. \end{array} \right.$$

En revanche, les paramètres μ_i et η_i , qui représentent respectivement le point d'équilibre d'un mouvement de type 3 et l'angle du drift d'un mouvement de type 2, sont propres à chaque trajectoire.

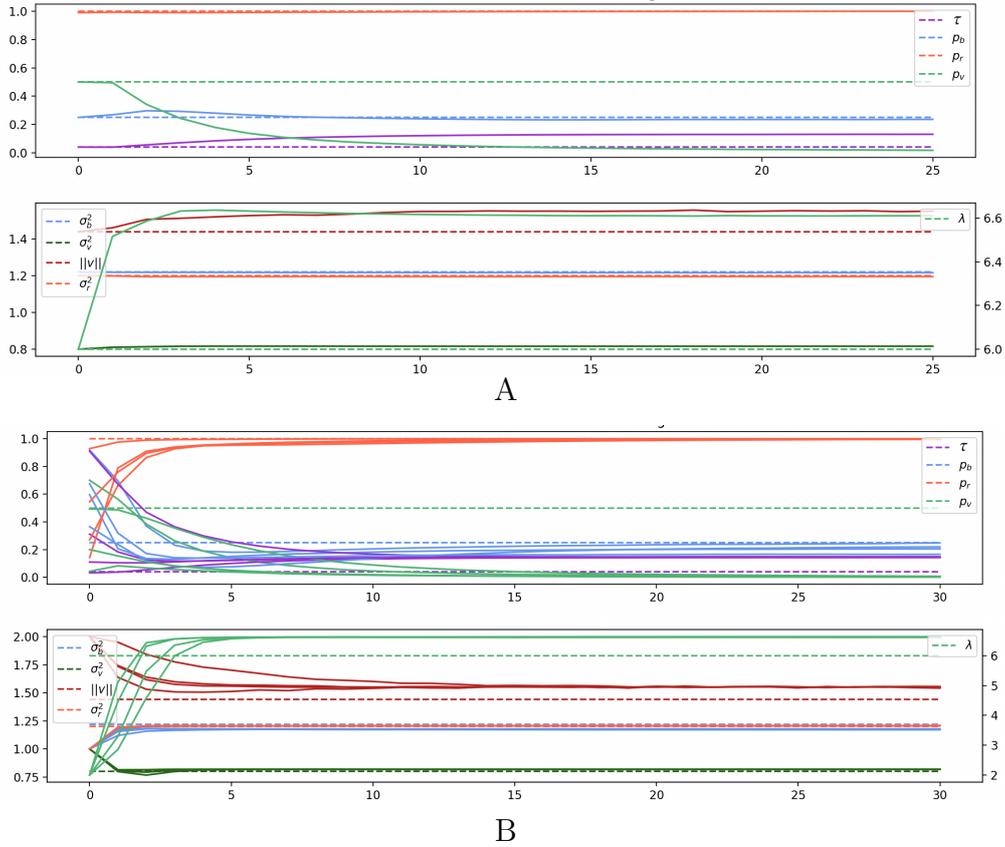


FIGURE 6.6 – Évolution du maximum de vraisemblance estimé par l’algorithme EM à chaque itération en fonction du nombre d’itérations. L’affichage de chaque sous-figure se fait en deux graphiques pour plus de lisibilité : en haut : graphique des paramètres τ , p_b , p_r , p_v et en bas : graphique des paramètres σ_b , ρ , σ_r , λ , σ_v . Ici, tous les paramètres sont supposés inconnus. Sous-figure A (haut) : les valeurs initiales des paramètres sont les vraies valeurs de chaque paramètre. Sous-figure B (bas) : plusieurs valeurs initiales différentes, choisies aléatoirement, ont été testées.

On utilise à présent la proposition 13 pour mettre en place l’algorithme EM. Par la suite, nous allons dans un premier temps considérer que tous les paramètres sont supposés inconnus, puis nous allons restreindre notre étude au cas où tous les paramètres sont inconnus, sauf les points d’équilibre μ_i et les angles η_i .

6.1.3.1 Cas où les points d’équilibres μ_i et les angles η_i sont inconnus

Dans cette section, on suppose inconnus tous les paramètres de l’étude, c’est-à-dire τ , p_b , p_r , p_v , σ_b , ρ , σ_r , λ , σ_v mais aussi μ_i et η_i pour tout $i \in \llbracket 1, s \rrbracket$. Dans ce cas là,

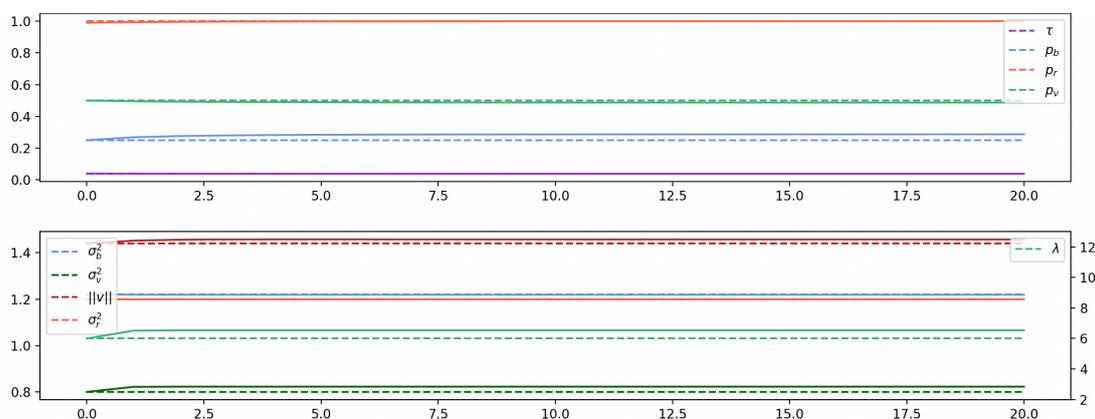


FIGURE 6.7 – Évolution du maximum de vraisemblance estimé par l’algorithme EM à chaque itération en fonction du nombre d’itérations. Ici, les angles et les points d’équilibres sont fixés à leur valeur théorique et ne sont donc pas estimés et la valeur initiale des paramètres estimés est leur valeur théorique. L’affichage est proposé sous forme de deux graphiques pour plus de lisibilité : en haut : graphique des paramètres τ , p_b , p_r , p_v et en bas : graphique des paramètres σ_b , ρ , σ_r , λ , σ_v (avec échelle à droite pour le paramètre λ).

la sous-figure A de la figure 6.6 illustre le fait que même lorsque l’on donne comme valeurs initiales à l’algorithme EM, les valeurs théoriques des différents paramètres, celui-ci ne converge pas. Il n’y a donc aucun espoir de faire converger cet algorithme, quelles que soient les valeurs initiales choisies aléatoirement.

Ce constat est confirmé si on exécute l’algorithme EM en estimant tous les paramètres (y compris les angles et les points d’équilibres, comme précédemment) pour plusieurs valeurs initiales aléatoires, comme le montre la sous-figure B de la figure 6.6.

On observe dans la sous-section suivante, illustrée par la figure 6.7, que l’algorithme converge dans le cas où l’expérience est réalisée en supposant les angles η_i et les points d’équilibres μ_i connus. Ainsi, au vu de ces deux résultats (figures 6.6 et 6.7), on peut en conclure que l’algorithme EM est trop sensible à la variation des points d’équilibre et des angles de chaque trajectoire.

6.1.3.2 Cas où les points d’équilibres μ_i et les angles η_i sont connus

Dans cette section, on suppose inconnus les paramètres σ_b , ρ , σ_r , λ , σ_v et τ , p_b , p_r , p_v . En revanche, contrairement à ce qui a été supposé dans le paragraphe précédent, les paramètres μ_i et η_i pour tout $i \in \llbracket 1, s \rrbracket$ sont connus. On exécute ensuite

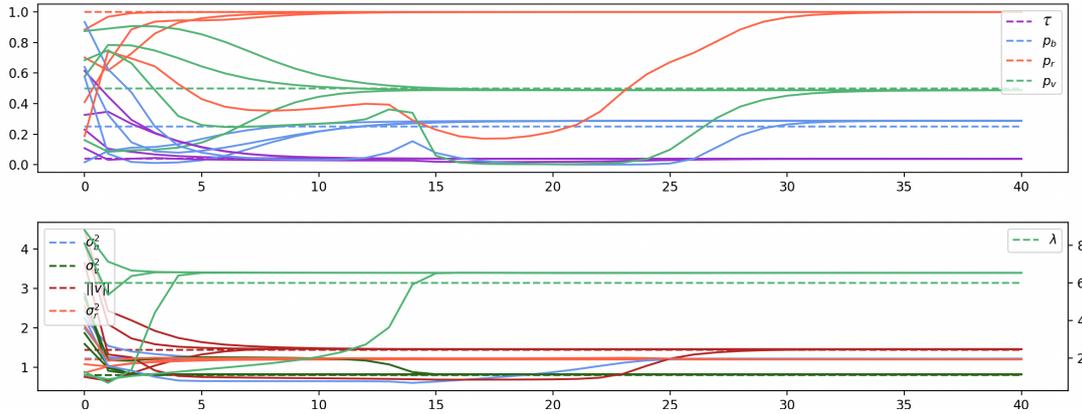


FIGURE 6.8 – Évolution du maximum de vraisemblance estimé par l’algorithme EM à chaque itération en fonction du nombre d’itérations. Ici, les angles et les points d’équilibre sont fixés égaux à leurs vraies valeurs et ne sont donc pas estimés et la valeur initiale des paramètres estimés est choisie uniformément dans $[0, 1]$ pour τ , p_b , p_r , p_v , uniformément dans $[0, 3]$ pour σ_b , σ_r , σ_v , uniformément dans $[0, 5]$ pour v et uniformément dans $[0, 10]$ pour λ . Quatre valeurs initiales ont été testées ici. L’affichage est toujours proposé sous forme de deux graphiques.

l’algorithme EM dans les mêmes conditions que dans la section 6.1.3.1, c’est-à-dire en prenant comme valeurs initiales les valeurs théoriques pour tous les paramètres mais en supposant cette fois-ci les paramètres μ_i et η_i pour tout $i \in \llbracket 1, s \rrbracket$ connus et égaux à leur valeur théorique. La figure 6.7 illustre la convergence de l’algorithme. La figure 6.8 met en évidence que l’algorithme EM converge toujours avec plusieurs valeurs initiales choisies aléatoirement.

6.1.3.3 Détermination des angles η_i et des points d’équilibre μ_i

Le fait d’étudier des données simulées permet de fixer les valeurs des angles et des points d’équilibre. Néanmoins, notre objectif est d’évaluer notre méthode sur un jeu de données réelles. On ne connaît pas les valeurs théoriques de l’angle et du point d’équilibre de chaque trajectoire. Il faut donc proposer une autre manière d’estimer ces deux quantités pour chaque trajectoire. Pour l’angle η_i , on choisit l’angle défini par la direction (par rapport à l’axe horizontal) du vecteur formé par le dernier point de la trajectoire et son premier point. Pour le point d’équilibre, il faut identifier un point “au milieu” de la partie verte d’une trajectoire. Pour trouver un tel point pour une trajectoire de longueur n , une idée est de chercher le cercle dit *maximum*, qui

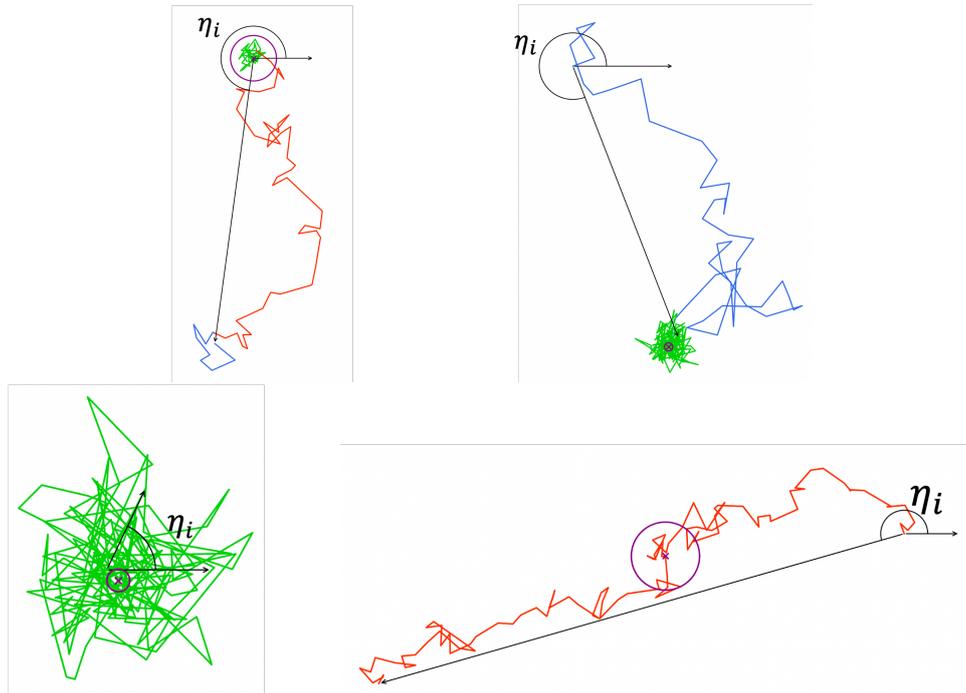


FIGURE 6.9 – Pour 4 trajectoires données, on représente le point d'équilibre (par la croix violette) choisi par la méthode du cercle maximum, en violet on retrouve aussi le cercle maximum associé. Cette figure représente aussi l'angle η_i défini par le vecteur formé par le premier et le dernier point de la trajectoire.

est un cercle de rayon :

$$r = \frac{k_x + k_y}{2} \quad \text{où} \quad k_x = \frac{IQ_x}{n^{1/3}} \quad \text{et} \quad k_y = \frac{IQ_y}{n^{1/3}},$$

avec IQ_x (resp. IQ_y) la distance inter-quantile de la liste formée des abscisses (resp. des ordonnées) des n points de la trajectoire. Pour chaque point x de la trajectoire, on compte le nombre n_x de points de la trajectoire présents dans le cercle de centre x et de rayon r et on appelle cercle *maximum* le cercle de centre x pour lequel ce nombre n_x est le plus grand. On choisit ensuite pour point d'équilibre le centre du cercle maximum.

La figure 6.9 affiche, pour 4 trajectoires données (issues de la figure 6.1), en violet le cercle maximum et avec une croix violette son centre (donc le point d'équilibre choisi), ainsi que l'angle η_i entre l'axe des abscisses et le vecteur formé par le premier et le dernier point de la trajectoire.

En utilisant l'algorithme EM dans les mêmes conditions que celles de la figure 6.8,

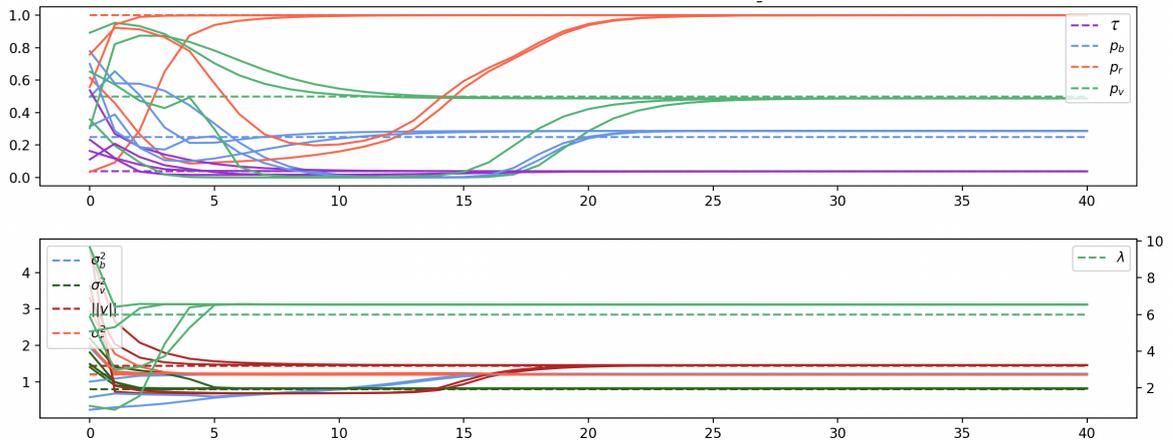


FIGURE 6.10 – Évolution du maximum de vraisemblance estimé par l'algorithme EM à chaque itération en fonction du nombre d'itérations. Ici, les angles et les points d'équilibres ne sont pas estimés par l'algorithme EM. Ils sont fixés par la méthode du cercle maximum pour le point d'équilibre et à l'angle défini par le vecteur formé par le premier et le dernier point de chaque trajectoire pour l'angle (cf section 6.1.3.3). La valeur initiale des paramètres estimés est choisie uniformément dans $[0, 1]$ pour τ , p_b , p_r , p_v , uniformément dans $[0, 3]$ pour σ_b , σ_r , σ_v , uniformément dans $[0, 5]$ pour v et uniformément dans $[0, 10]$ pour λ . Cinq valeurs initiales ont été testées ici. L'affichage est toujours proposé sous forme de deux graphiques.

c'est-à-dire en considérant les points d'équilibre et les angles de chaque trajectoire connus mais cette fois en leur donnant comme valeur respectivement le centre du cercle maximum et l'angle défini par le vecteur formé du premier et du dernier point de la trajectoire, on obtient le graphique de la figure 6.10, qui illustre la convergence de l'algorithme converge toujours très bien.

6.1.4 Bilan de l'étude par simulation

Nous faisons un rapide bilan de l'étude par simulation, en parcourant les différents scénarios étudiés.

Bilan 1. Si tous les paramètres du modèle sont considérés comme des inconnues pour l'algorithme EM, alors ce dernier ne converge pas correctement, même en prenant comme valeur initiale la valeur théorique de chaque paramètre, ou toutes autres valeurs choisies aléatoirement.

Il est donc nécessaire de retirer certains paramètres du vecteur des inconnues de l'algorithme EM, comme décrit ci-dessous.

Bilan 2. Si les seuls paramètres inconnus sont ceux de la matrice de transition a , c'est-à-dire τ , p_b , p_r et p_v et que tous les autres sont fixés avec les valeurs théoriques, alors l'algorithme **converge**, quelles que soient les valeurs initiales choisies aléatoirement.

Bilan 3. Si les seuls paramètres inconnus sont ceux d'une trajectoire i , c'est-à-dire $\sigma_{i,1}$, ρ_i , $\sigma_{i,2}$, λ_i , $\sigma_{i,3}$, η_i et μ_i alors l'algorithme **ne converge pas toujours**. Notre méthode n'est donc pas adaptée pour étudier une trajectoire individuelle. On l'utilisera pour l'étude d'un grand nombre de trajectoires.

Il y a un intérêt à étudier un grand nombre de trajectoires seulement si ces dernières suivent trois mouvements distincts ayant les mêmes caractéristiques. C'est pourquoi, nous supposons par la suite que pour tout $i \in \llbracket 1, s \rrbracket$, $\sigma_{i,1} = \sigma_b$, $\rho_i = \rho$, $\sigma_{i,2} = \sigma_r$, $\lambda_i = \lambda$, $\sigma_{i,3} = \sigma_v$.

Sous cette hypothèse, les résultats obtenus permettent de conclure :

Bilan 4. Si tous les paramètres du modèle sont considérés comme des inconnues pour l'algorithme EM sauf les angles η_i et les points d'équilibres μ_i , alors l'algorithme **converge** si les angles et les points d'équilibres sont fixés à leurs valeurs théoriques, quelles que soient les valeurs initiales.

Sur la base des bilans 1 et 4, on déduit que l'algorithme converge si les angles et les points d'équilibres ne sont pas estimés par l'algorithme. Il faut donc les estimer au préalable.

Bilan 5. Si tous les paramètres du modèle sont considérés comme des inconnues pour l'algorithme EM sauf les angles η_i et les points d'équilibres μ_i , qui sont respectivement fixés par la méthode du cercle maximum pour le point d'équilibre et par l'angle défini par le vecteur formé par le premier et le dernier point de chaque trajectoire pour l'angle (cf section 6.1.3.3), alors l'algorithme **converge**, quelles que soient les valeurs initiales choisies.

Le bilan 5 conduit à une méthode qu'il est possible de mettre en place pour estimer tous les paramètres des 3 types de mouvements (σ_b , ρ , σ_v , λ , σ_v) lorsque nous ne connaissons pas les angles η_i et les points d'équilibre μ_i . C'est cette méthode que nous utiliserons dans la suite pour analyser les données réelles.

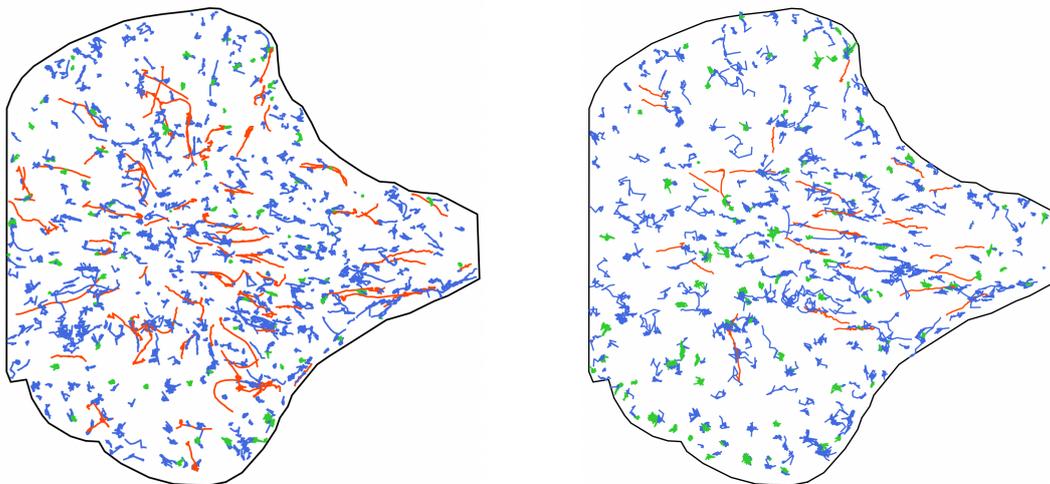


FIGURE 6.11 – Trajectoires des protéines Langerin (à gauche) et Rab-11 (à droite) classées avec la méthode de [BVK19] avec toujours le même code couleur : **bleu** pour les morceaux de trajectoires browniens, **rouge** pour ceux superdiffusifs et **vert** pour ceux sousdiffusifs.

6.2 Résultats sur des données réelles

Dans cette section, nous évaluons la performance de l'algorithme EM sur des données réelles et nous comparons les résultats avec ceux obtenus par d'autres méthodes déjà existantes.

Les données réelles étudiées dans ce paragraphes sont les données qui ont été présentées aux chapitres 2 et 4 : il s'agit des protéines Langerin et Rab-11, qui se déplacent à la surface d'une cellule. Ces données ont été étudiées par Vincent Briane dans [BVK19]. Sa méthode calcule une statistique de test (basée sur l'excursion maximale par rapport au premier point de temps) sur une fenêtre glissante sur chaque trajectoire pour détecter les changements de régime. La figure 6.11 affiche les résultats obtenus par cette méthode sur le jeu de données réelles.

Nous testons l'algorithme EM sur ces données en se plaçant dans les mêmes conditions que celles de la section 6.1.3.2, c'est-à-dire en supposant tous les paramètres inconnus sauf les points d'équilibres et les angles de chaque trajectoire qui sont estimés en amont grâce à la méthode du cercle maximum et de l'angle formé par le vecteur reliant le dernier et le premier point de la trajectoire (cf section 6.1.3.3). On suppose aussi que les trois mouvements possibles (1, 2 et 3) possèdent le même

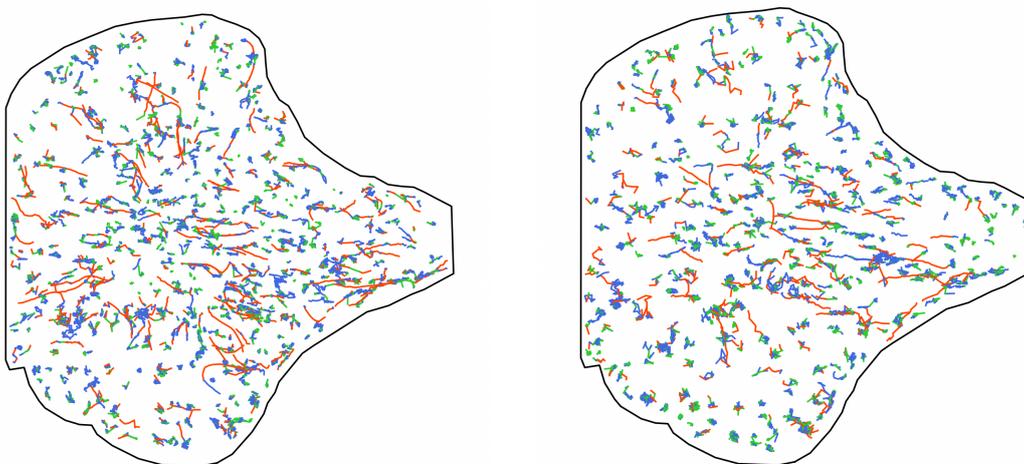


FIGURE 6.12 – Trajectoires des protéines Langerin (à gauche) et Rab-11 (à droite) classées avec notre méthode avec toujours le même code couleur : **bleu** pour les morceaux de trajectoires browniens, **rouge** pour ceux superdiffusifs et **vert** pour ceux sousdiffusifs.

coefficient de diffusion σ , c'est-à-dire que $\sigma_b = \sigma_r = \sigma_v = \sigma$ (on se rapporte donc à la proposition 14 pour le calcul de l'argument du maximum de Q). La figure 6.13 illustre l'évolution de la convergence de l'algorithme au cours des itérations. La figure 6.12 affiche les trajectoires classées grâce à l'algorithme de Viterbi, en se servant des résultats de l'algorithme EM sur ces données (ceux de la figure 6.13).

Comparaison des résultats des figures 6.11 et 6.12 : Les résultats obtenus avec la méthode de [BVK19] présentent beaucoup plus de trajectoires de type 1. Ce phénomène est amplifié par le fait que cette méthode classe automatiquement en type 1 les trajectoires trop courtes. De plus, les trajectoires de la figure 6.11 présentent très peu de changements de régime : une quinzaine pour les protéines Langerin (gauche) contre environ 2300 pour les mêmes trajectoires de la figure 6.12. Une telle différence peut être expliquée par le fait que dans la méthode de [BVK19], la présence ou non de changement de régime est directement liée à la taille de la fenêtre glissante choisie par l'utilisateur. De plus, la méthode de [BVK19] traite chaque trajectoire individuellement et ne tient pas compte d'effet de population. Visuellement, la figure 6.12 s'appuyant sur notre méthode, présente beaucoup plus de trajectoires confinées, classées en type 3 que la figure 6.11 et de trajectoires dirigées classées en type 2.

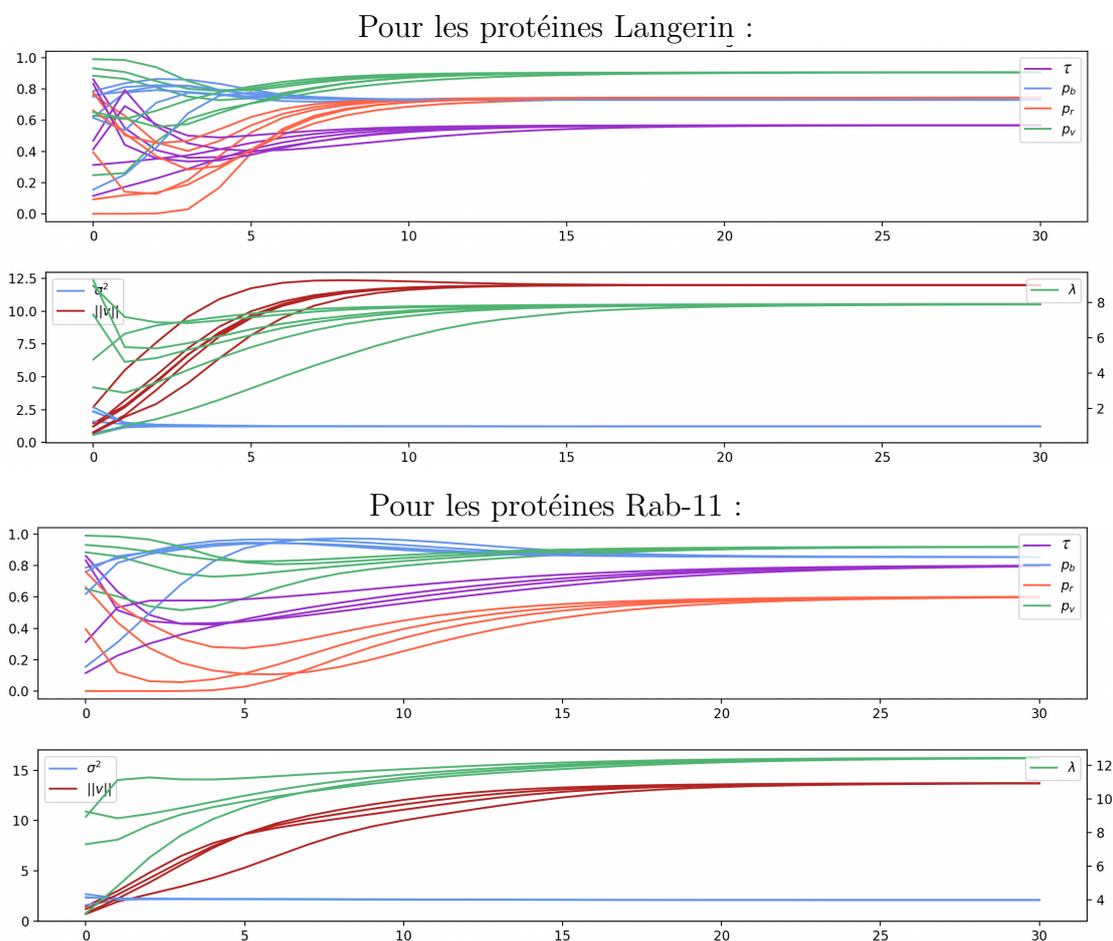


FIGURE 6.13 – Évolution du maximum de vraisemblance estimé par l’algorithme EM à chaque itération en fonction du nombre d’itérations. Ici, les angles et les points d’équilibre sont supposés connus et ne sont donc pas estimés. Ils sont fixés par la méthode du cercle maximum pour le point d’équilibre et à l’angle défini par le vecteur formé par le premier et le dernier point de chaque trajectoire pour l’angle (cf section 6.1.3.3). La valeur initiale des paramètres estimés est choisie aléatoirement dans $[0, 1]$ pour τ , p_b , p_r , p_v , aléatoirement dans $[0, 3]$ pour σ , aléatoirement dans $[0, 5]$ pour v et aléatoirement dans $[0, 10]$ pour λ . Quatre valeurs initiales ont été testées ici. L’affichage est toujours proposé sous forme de deux graphiques.

6.3 Conclusion

Dans ce chapitre, nous avons présenté les résultats de notre méthode de classification (présentée au chapitre précédent) appliquée sur un jeu de données simulées.

Les résultats sont très satisfaisants dans le cas où les paramètres sont communs à toutes les trajectoires (hypothèse d’homogénéité en population). En revanche, l’estimation menée sur une seule trajectoire est moins fiable. Il est recommandé, dans ce cas, de privilégier d’autres méthodes basées sur des principes différents (comme celle de [MBP⁺15], présentée en annexe H). Un bilan global des résultats est reporté dans la section 6.1.4. Pour conclure ce chapitre, nous avons classé les trajectoires issues du jeu de données réelles commun à tous les chapitres du manuscrit. Les résultats sont assez satisfaisants d’un point de vue qualitatif, mais plus difficiles à évaluer d’un point de vue quantitatif sans vérité-terrain.

Annexes de la partie III

FORME DE LA FONCTION Q DANS LE CAS DE TRAJECTOIRES INDÉPENDANTES

Le but de cette annexe est de prouver le résultat de la proposition 6 du chapitre 5, qui est aussi la proposition 4 de cette annexe.

Ainsi, nous allons démontrer que si les conditions **(C1)**, **(C2)** et **(C3)**, que nous rappelons ici, sont satisfaites, alors la fonction Q se calcule comme si les mêmes trajectoires étaient considérées sous forme de “stacking”.

- (C1)** L'intensité de saut α est bornée.
- (C2)** Le paramètre d'intérêt ϑ n'intervient que dans τ , k_τ et L .
- (C3)** Les mouvements des trajectoires et leurs changements de régime sont indépendants au sens de la définition 1 de cette annexe.

F.1 Notations

Dans toute cette annexe, on prendra soin de noter en gras les éléments qui sont des vecteurs et en lettres normales les éléments qui sont des valeurs réelles. Ainsi, les coordonnées d'un vecteur \mathbf{z} s'écrivent z_1, \dots, z_n .

Dans l'expression $Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \mathbb{E}_{\tilde{\vartheta}} \left(\ln(p(\mathbf{Z}, \mathbf{M}; \tilde{\vartheta}) | \mathbf{Z}) \right)$, \mathbf{Z} désigne toutes les positions du processus observé entre t_0 et t_n , que l'on a choisi de regrouper trajectoire par trajectoire dans le chapitre 5 :

$$\mathbf{Z} = (\mathbf{Z}^1, \dots, \mathbf{Z}^s),$$

où $\mathbf{Z}^i = (z_{t_b,i}, \dots, z_{t_d,i})$ est l'ensemble des positions de la trajectoire i .

Dans cette annexe, nous allons désormais réécrire \mathbf{Z} de la façon suivante :

$$\mathbf{Z} = (\mathbf{z}_{t_0}, \dots, \mathbf{z}_{t_n}),$$

où les $(t_i)_{0 \leq i \leq n}$ sont les temps discrets d'observation du processus et où \mathbf{z}_{t_i} représente l'ensemble des positions des trajectoires présentes au temps t_i . Tous les vecteurs \mathbf{z}_{t_i} sont de taille s et on trouve à la k -ième coordonnée de \mathbf{z}_{t_i} , notée $z_{t_i,k}$, la position de la trajectoire numéro k si celle-ci est présente au temps t_i et le symbole \emptyset si celle-ci n'est pas présente au temps t_i . Ainsi $z_{t_i,k} \in \Lambda_\emptyset := \Lambda \cup \{\emptyset\}$.

De la même manière, l'ensemble des marques \mathbb{M} sera modifié en

$$\mathbb{M}_\emptyset := \mathbb{M} \cup \{\emptyset\} = \llbracket 1, M \rrbracket \cup \{\emptyset\}$$

et le vecteur \mathbf{M} des états cachés sera aussi modifié selon le même point de vue :

$$\mathbf{M} = (\mathbf{m}_{t_0}, \dots, \mathbf{m}_{t_n}),$$

où les vecteurs \mathbf{m}_{t_i} seront désormais à valeurs dans $(\mathbb{M}_\emptyset)^s$ et plus précisément dans :

$$\mathbb{M}_\emptyset^{(\mathbf{z}_{t_i})} := \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_s \end{pmatrix}, a_j \in \llbracket 1, M \rrbracket \cup \{\emptyset\} \text{ et } a_j = \emptyset \text{ ssi } z_{t_i,j} = \emptyset \right\}.$$

Dans la suite, on notera pour tout $\mathbf{z} \in (\Lambda_\emptyset)^s$, $\mathcal{P}(\mathbf{z}) = \{j \in \llbracket 1, s \rrbracket, z_j \neq \emptyset\}$ l'ensemble des numéros des trajectoires présentes dans la configuration \mathbf{z} .

On introduit trois nouveaux ensembles : pour $A, B \in \{\Lambda, \mathbb{M}\}$, on note :

$$\begin{aligned} \mathcal{I}_1^{A \times B} &= \{(\mathbf{y}, \mathbf{z}) \in A^s \times B^s, \forall i \in \llbracket 1, s \rrbracket, y_i = \emptyset \Leftrightarrow z_i = \emptyset\}, \\ \mathcal{I}_2^{A \times B} &= \{(\mathbf{y}, \mathbf{z}) \in A^s \times B^s, \exists ! i_0 \in \llbracket 1, s \rrbracket / \{i \in \llbracket 1, s \rrbracket / y_i = \emptyset \text{ et } z_i \neq \emptyset\} = \{i_0\} \\ &\quad \text{et } \forall i \in \llbracket 1, s \rrbracket \setminus \{i_0\}, y_i = \emptyset \Leftrightarrow z_i = \emptyset\}, \\ \mathcal{I}_3^{A \times B} &= \{(\mathbf{y}, \mathbf{z}) \in A^s \times B^s, \exists ! i_0 \in \llbracket 1, s \rrbracket / \{i \in \llbracket 1, s \rrbracket / z_i = \emptyset \text{ et } y_i \neq \emptyset\} = \{i_0\} \\ &\quad \text{et } \forall i \in \llbracket 1, s \rrbracket \setminus \{i_0\}, y_i = \emptyset \Leftrightarrow z_i = \emptyset\}. \end{aligned}$$

Si $A = B$, on notera $\mathcal{I}_j^{A \times A} =: \mathcal{I}_j^A$ pour $j \in \{1, 2, 3\}$. On note la réunion des trois ensembles :

$$\mathcal{I}^\Lambda := \mathcal{I}_1^\Lambda \cup \mathcal{I}_2^\Lambda \cup \mathcal{I}_3^\Lambda \quad \text{et} \quad \mathcal{I}^\mathbb{M} := \mathcal{I}_1^\mathbb{M} \cup \mathcal{I}_2^\mathbb{M} \cup \mathcal{I}_3^\mathbb{M}.$$

Dans \mathcal{I}_1^Λ , on retrouve les vecteurs (\mathbf{y}, \mathbf{z}) qui comportent les mêmes trajectoires présentes. Ainsi si aucune naissance, ni aucune mort n'a lieu entre deux configurations \mathbf{y} et \mathbf{z} , alors $(\mathbf{y}, \mathbf{z}) \in \mathcal{I}_1^\Lambda$. Dans \mathcal{I}_2^Λ , on retrouve les vecteurs (\mathbf{y}, \mathbf{z}) pour lesquels la configuration \mathbf{z} comporte une trajectoire présente de plus que \mathbf{y} . Ainsi, si une seule naissance et aucune mort a lieu entre les configurations \mathbf{y} et \mathbf{z} , alors $(\mathbf{y}, \mathbf{z}) \in \mathcal{I}_2^\Lambda$. Dans \mathcal{I}_3^Λ , on retrouve les vecteurs (\mathbf{y}, \mathbf{z}) pour lesquels la configuration \mathbf{y} comporte une trajectoire présente de plus que \mathbf{z} . Ainsi, si une seule mort et aucune naissance a lieu entre les configurations \mathbf{y} et \mathbf{z} , alors $(\mathbf{y}, \mathbf{z}) \in \mathcal{I}_3^\Lambda$.

De plus, si $(\mathbf{y}, \mathbf{z}) \in \mathcal{I}_2^\Lambda$, alors par définition, il existe un unique $i \in \llbracket 1, s \rrbracket$, que l'on note i_0 tel que $y_{i_0} = \emptyset$ et $z_{i_0} \neq \emptyset$, on note alors $\mathbf{z}^{(\mathbf{y})}$ le vecteur de coordonnées :

$$\forall i \in \llbracket 1, s \rrbracket \setminus \{i_0\}, (z^{(\mathbf{y})})_i = z_i \text{ et } (z^{(\mathbf{y})})_{i_0} = \emptyset.$$

De la même manière, si $(\mathbf{y}, \mathbf{z}) \in \mathcal{I}_3$, alors par définition, il existe un unique $i \in \llbracket 1, s \rrbracket$, que l'on note i_0 tel que $z_{i_0} = \emptyset$ et $y_{i_0} \neq \emptyset$, on note alors $\mathbf{y}^{(\mathbf{z})}$ le vecteur de coordonnées :

$$\forall i \in \llbracket 1, s \rrbracket \setminus \{i_0\}, (y^{(\mathbf{z})})_i = y_i \text{ et } (y^{(\mathbf{z})})_{i_0} = \emptyset.$$

Ainsi, si $(\mathbf{y}, \mathbf{z}) \in \mathcal{I}_2^\Lambda$, on a $\mathcal{P}(\mathbf{z}^{(\mathbf{y})}) = \mathcal{P}(\mathbf{y})$. De même, si $(\mathbf{y}, \mathbf{z}) \in \mathcal{I}_3^\Lambda$, on a $\mathcal{P}(\mathbf{y}^{(\mathbf{z})}) = \mathcal{P}(\mathbf{z})$.

F.2 Cadre

Comme le processus BDMM continu $(\mathbf{Z}_t, \mathbf{M}_t)_{t \geq 0}$, duquel sont extraites les trajectoires discrétisées est un processus de Markov homogène [LLG21], les quatre hypothèses suivantes sont vérifiées :

(H1) Pour tout $i \in \{1, \dots, n\}$, \mathbf{m}_{t_i} est indépendante de toutes les autres variables

conditionnellement à $\mathbf{z}_{t_{i-1}}$ et $\mathbf{m}_{t_{i-1}}$, i.e. :

$$p\left(\mathbf{m}_{t_i} | \mathbf{m}_{t_{i-1}}, \dots, \mathbf{m}_{t_0}, \mathbf{z}_{t_{i-1}}, \dots, \mathbf{z}_{t_0}\right) = p\left(\mathbf{m}_{t_i} | \mathbf{m}_{t_{i-1}}, \mathbf{z}_{t_{i-1}}\right).$$

(H2) Pour tout $i \in \{1, \dots, n\}$, pour tout $\mathbf{m} \in (\mathbb{M}_\emptyset)^s$, pour tout $\mathbf{z} \in (\Lambda_\emptyset)^s$ et pour tout $\tilde{\mathbf{m}} \in \mathbb{M}_\emptyset^{(\mathbf{z})}$:

$$\mathbb{P}\left(\mathbf{m}_{t_i} = \mathbf{m} | \mathbf{m}_{t_{i-1}} = \tilde{\mathbf{m}}, \mathbf{z}_{t_{i-1}} = \mathbf{z}\right) = \mathbb{P}\left(\mathbf{m}_{t_1} = \mathbf{m} | \mathbf{m}_{t_0} = \tilde{\mathbf{m}}, \mathbf{z}_{t_0} = \mathbf{z}\right).$$

(H3) Pour tout $i \in \{1, \dots, n\}$, la position \mathbf{z}_{t_i} est indépendante de toutes les autres variables conditionnellement à \mathbf{m}_{t_i} et $\mathbf{z}_{t_{i-1}}$, i.e. :

$$p\left(\mathbf{z}_{t_i} | \mathbf{m}_{t_i}, \dots, \mathbf{m}_{t_0}, \mathbf{z}_{t_{i-1}}, \dots, \mathbf{z}_{t_0}\right) = p\left(\mathbf{z}_{t_i} | \mathbf{z}_{t_{i-1}}, \mathbf{m}_{t_i}\right).$$

(H4) Pour tout $i \in \{1, \dots, n\}$, pour tout $\tilde{\mathbf{z}} \in (\Lambda_\emptyset)^s$ et pour tout $\mathbf{m} \in (\mathbb{M}_\emptyset)^s$:

$$p\left(\mathbf{z}_{t_i} | \mathbf{z}_{t_{i-1}} = \tilde{\mathbf{z}}, \mathbf{m}_{t_i} = \mathbf{m}\right) = p\left(\mathbf{z}_{t_1} | \mathbf{z}_{t_0} = \tilde{\mathbf{z}}, \mathbf{m}_{t_1} = \mathbf{m}\right).$$

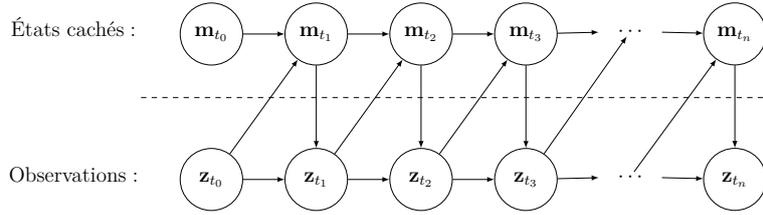
On introduit alors les trois quantités suivantes :

- La loi de probabilité des changements d'états, notée $A = \{A_{\mathbf{u}, \mathbf{v}}(\mathbf{z})\}$. Pour tout $i \in \llbracket 1, n \rrbracket$, pour tout $\mathbf{v} \in (\mathbb{M}_\emptyset)^s$, pour tout $\mathbf{z} \in (\Lambda_\emptyset)^s$ et pour tout $\mathbf{u} \in \mathbb{M}_\emptyset^{(\mathbf{z})}$, on a :

$$A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}) = \mathbb{P}\left(\mathbf{m}_{t_i} = \mathbf{v} | \mathbf{z}_{t_{i-1}} = \mathbf{z}, \mathbf{m}_{t_{i-1}} = \mathbf{u}\right).$$

- La loi de probabilité des observations, notée B . Pour $\mathbf{y} \in (\Lambda_\emptyset)^s$ et $\mathbf{u} \in (\mathbb{M}_\emptyset)^s$, la densité de \mathbf{z}_{t_i} sachant $\mathbf{z}_{t_{i-1}} = \mathbf{y}$ et $\mathbf{m}_{t_i} = \mathbf{v}$ est donnée par la fonction $\mathbf{z} \mapsto B(\mathbf{z} | \mathbf{y}, \mathbf{v})$.
- La loi initiale du couple $(\mathbf{z}_{t_0}, \mathbf{m}_{t_0})$ notée $\pi : (\Lambda_\emptyset \times \mathbb{M}_\emptyset)^s \rightarrow \mathbb{R}_+$.

Ainsi on a le schéma suivant :



Comme dans le chapitre 5, on notera $A_{\mathbf{u},\mathbf{v}}(\mathbf{z}; \vartheta)$ au lieu de $A_{\mathbf{u},\mathbf{v}}(\mathbf{z})$ et $B(\mathbf{z}|\mathbf{y}, \mathbf{v}; \vartheta)$ au lieu de $B(\mathbf{z}|\mathbf{y}, \mathbf{v})$ pour marquer la présence du paramètre ϑ dans ces fonctions.

Donnons maintenant la définition d'indépendance des mouvements des trajectoires et de leurs changements de régimes, notion que l'on supposera être satisfaite dans le chapitre 5.

Cette définition permet d'exprimer de manière mathématique et rigoureuse les deux notions très intuitives suivantes :

- *l'indépendance des mouvements des trajectoires*, qui va traduire le fait que le mouvement d'une particule dépend uniquement de la marque de la particule, et est indépendant du mouvement des autres particules alentours,
- *l'indépendance des changements de régime*, qui va traduire le fait que lorsqu'une mutation a lieu, alors ce changement de marque se fait indépendamment des positions et des marques des autres particules alentours.

Définition 1. On dit d'un processus BDMM que :

1. ses changements de régime sont indépendants si : pour tout $(\mathbf{u}, \mathbf{v}) \in \mathcal{I}^{\mathbb{M}}$, pour tout $\mathbf{z} \in (\Lambda_{\emptyset})^s$ tel que pour tout $j \in \llbracket 1, s \rrbracket$, $z_j = \emptyset$ si $u_j = \emptyset$, on a :

$$A_{\mathbf{u},\mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) \underset{\Delta \rightarrow 0}{\sim} C(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) \prod_{\substack{j=1 \\ j \in \mathcal{P}(\mathbf{u}) \cap \mathcal{P}(\mathbf{v})}}^s a_{u_j, v_j}(z_j; \tilde{\vartheta}),$$

où $C(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta)$ est une constante qui ne dépend pas de $\tilde{\vartheta}$ et où $a_{u,v}(z; \tilde{\vartheta})$ est défini dans l'équation (5.1) (section 5.2.1 du chapitre 5) par

$$a_{u,v}(z; \tilde{\vartheta}) = \mathbb{P} \left(m_{t_i, j} = v \mid z_{t_{i-1}, j} = z, m_{t_{i-1}, j} = u, \tilde{\vartheta} \right),$$

pour tout $j \in \llbracket 1, s \rrbracket$ et pour tout $i \in \llbracket j_b + 1, j_d \rrbracket$ (où j_b et j_d sont aussi définis en section 5.2.1).

2. les mouvements de ses trajectoires sont indépendants si : pour tout $(\mathbf{y}, \mathbf{z}) \in \mathcal{I}^\Lambda$, on a pour tout $\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z})}$:

$$L(\mathbf{z}|\mathbf{y}, \mathbf{v}; \tilde{\vartheta}) = \prod_{\substack{j=1 \\ j \in \mathcal{P}(\mathbf{z}) \cap \mathcal{P}(\mathbf{y})}}^s b^j(z_j|y_j, v_j; \tilde{\vartheta}),$$

où $y \mapsto b^j(y|z, v; \tilde{\vartheta})$ est défini dans l'équation (5.2) (section 5.2.1 du chapitre 5) comme la vraisemblance de $z_{t_k, j}$ sachant que $z_{t_{k-1}, j} = z \in \Lambda_{\mathcal{O}}$ et $m_{t_k, j} = v \in \mathbb{M}_{\mathcal{O}}$, pour tout $k \in \llbracket j_b + 1, j_d \rrbracket$.

F.3 Calcul de la fonction Q

Dans cette section, nous allons calculer, pour le processus $(\mathbf{Z}_t, \mathbf{M}_t)_t$ considéré, la fonction : $Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \mathbb{E}_{\vartheta}(\ln(p(\mathbf{Z}, \mathbf{M}; \tilde{\vartheta})|\mathbf{Z}))$. Pour cela, nous allons d'abord détailler la forme de cette fonction Q dans le cas où le processus BDMM considéré est observé à des temps discrets t_0, \dots, t_n . Dans un deuxième temps, nous allons simplifier cette expression générale pour traiter le cas où le paramètre d'intérêt ϑ n'intervient que dans les parties "move" (fonction L) et mutation (fonctions τ et k_{τ}). Pour finir, nous simplifierons encore la forme de Q dans le cas où on ajoute l'hypothèse d'indépendance des mouvements des trajectoires et de leurs changements de régime.

F.3.1 Forme de Q dans le cas d'un BDMM observé en temps discret

Calculons maintenant la fonction Q dans le cas où le processus BDMM est observé à des temps discrets t_0, \dots, t_n .

On rappelle que :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \mathbb{E}_{\vartheta}(\ln(p(\mathbf{Z}, \mathbf{M}; \tilde{\vartheta})|\mathbf{Z})) = \mathbb{E}_{\vartheta}(\ln(p(\mathbf{Z}_{t_0:t_n}, \mathbf{M}_{t_0:t_n}; \tilde{\vartheta})|\mathbf{Z}_{t_0:t_n})).$$

Proposition 1. On a :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \sum_{k=1}^n \sum_{\mathbf{u} \in \mathbb{M}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v}, \mathbf{m}_{t_{k-1}} = \mathbf{u} | \mathbf{Z}; \vartheta) \ln(A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \\ + \sum_{k=1}^n \sum_{\mathbf{v} \in \mathbb{M}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}; \vartheta) \ln(B(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta})) + C(\vartheta, \mathbf{Z}, \Delta),$$

où $C(\vartheta, \mathbf{Z}, \Delta)$ ne dépend pas de $\tilde{\vartheta}$.

Démonstration. On a :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \sum_{\mathbf{m} \in \bigotimes_{i=0}^n \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_i})}} p(\mathbf{M}_{t_0:t_n} = \mathbf{m} | \mathbf{Z}, \vartheta) \ln(p(\mathbf{Z}, \mathbf{m}; \tilde{\vartheta})).$$

Par une succession d'utilisations de la formule des probabilités composées, pour $\mathbf{m} = (\mathbf{m}_0, \dots, \mathbf{m}_n) \in \bigotimes_{i=0}^n \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_i})}$, on a :

$$p(\mathbf{M}_{t_0:t_n} = \mathbf{m}, \mathbf{Z}; \tilde{\vartheta}) \\ = p(\mathbf{M}_{t_0:t_{n-1}} = \mathbf{m}_{0:n-1}, \mathbf{Z}_{t_0:t_{n-1}}; \tilde{\vartheta}) \\ \times p(z_{t_n} | \mathbf{M}_{t_0:t_n} = \mathbf{m}_{0:n}, \mathbf{Z}_{t_0:t_{n-1}}; \tilde{\vartheta}) p(\mathbf{m}_{t_n} = \mathbf{m}_n | \mathbf{M}_{t_0:t_{n-1}} = \mathbf{m}_{0:n-1}, \mathbf{Z}_{t_0:t_{n-1}}; \tilde{\vartheta}) \\ = p(\mathbf{m}_{t_0} = \mathbf{m}_0, \mathbf{Z}_{t_0:t_0}; \tilde{\vartheta}) \\ \times \prod_{j=1}^n p(z_{t_j} | \mathbf{M}_{t_0:t_j} = \mathbf{m}_{0:j}, \mathbf{Z}_{t_0:t_{j-1}}; \tilde{\vartheta}) p(\mathbf{m}_{t_j} = \mathbf{m}_j | \mathbf{M}_{t_0:t_{j-1}} = \mathbf{m}_{0:j-1}, \mathbf{Z}_{t_0:t_{j-1}}; \tilde{\vartheta}).$$

D'une part, $p(\mathbf{m}_{t_0} = \mathbf{m}_0, \mathbf{Z}_{t_0:t_0}; \tilde{\vartheta}) = p(\mathbf{z}_{t_0}, \mathbf{m}_{t_0} = \mathbf{m}_0; \tilde{\vartheta}) = \pi(\mathbf{z}_{t_0}, \mathbf{m}_0)$, et d'autre part pour $j \in \llbracket 1, n \rrbracket$:

$$p(\mathbf{z}_{t_j} | \mathbf{M}_{t_0:t_j} = \mathbf{m}_{0:j}, \mathbf{Z}_{t_0:t_{j-1}}; \tilde{\vartheta}) p(\mathbf{m}_{t_j} = \mathbf{m}_j | \mathbf{M}_{t_0:t_{j-1}} = \mathbf{m}_{0:j-1}, \mathbf{Z}_{t_0:t_{j-1}}; \tilde{\vartheta}) \\ = p(\mathbf{z}_{t_j} | \mathbf{z}_{t_{j-1}}, \mathbf{m}_{t_j} = \mathbf{m}_j; \tilde{\vartheta}) p(\mathbf{m}_{t_j} = \mathbf{m}_j | \mathbf{m}_{t_{j-1}} = \mathbf{m}_{j-1}, \mathbf{z}_{t_{j-1}}; \tilde{\vartheta}) \\ = B(\mathbf{z}_{t_j} | \mathbf{z}_{t_{j-1}}, \mathbf{m}_j; \tilde{\vartheta}) A_{\mathbf{m}_{j-1}, \mathbf{m}_j}(\mathbf{z}_{t_{j-1}}; \tilde{\vartheta}),$$

donc :

$$\ln(p(\mathbf{M}_{t_0:t_n} = \mathbf{m}, \mathbf{Z}; \tilde{\vartheta})) = \ln(\pi(\mathbf{z}_{t_0}, \mathbf{m}_0)) \\ + \sum_{k=1}^n (\ln(A_{\mathbf{m}_{k-1}, \mathbf{m}_k}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) + \ln(B(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{m}_k; \tilde{\vartheta}))).$$

Ainsi,

$$\begin{aligned}
 Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = & C(\vartheta, \mathbf{Z}, \Delta) + \sum_{\mathbf{m} \in \bigotimes_{i=0}^n \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_i})}} \mathbb{P}(\mathbf{M}_{t_0:t_n} = \mathbf{m} | \mathbf{Z}; \vartheta) \sum_{k=1}^n \ln(A_{\mathbf{m}_{k-1}, \mathbf{m}_k}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \\
 & + \sum_{\mathbf{m} \in \bigotimes_{i=0}^n \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_i})}} \mathbb{P}(\mathbf{M}_{t_0:t_n} = \mathbf{m} | \mathbf{Z}; \vartheta) \sum_{k=1}^n \ln(B(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{m}_k; \tilde{\vartheta})), \quad (\text{F.1})
 \end{aligned}$$

où :

$$C(\vartheta, \mathbf{Z}, \Delta) = \sum_{\mathbf{m} \in \bigotimes_{i=0}^n \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_i})}} \mathbb{P}(\mathbf{M}_{t_0:t_n} = \mathbf{m} | \mathbf{Z}; \vartheta) \ln(\pi(\mathbf{z}_{t_0}, \mathbf{m}_0)).$$

Pour le deuxième terme de la somme (F.1), on a :

$$\begin{aligned}
 & \sum_{\mathbf{m} \in \bigotimes_{i=0}^n \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_i})}} \mathbb{P}(\mathbf{M}_{t_0:t_n} = \mathbf{m} | \mathbf{Z}; \vartheta) \sum_{k=1}^n \ln(A_{\mathbf{m}_{k-1}, \mathbf{m}_k}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \\
 = & \sum_{k=1}^n \sum_{\mathbf{m}_0 \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_0})}} \dots \sum_{\mathbf{m}_n \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_n})}} \mathbb{P}(\mathbf{m}_{t_0} = \mathbf{m}_0, \dots, \mathbf{m}_{t_n} = \mathbf{m}_n | \mathbf{Z}; \vartheta) \ln(A_{\mathbf{m}_{k-1}, \mathbf{m}_k}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \\
 = & \sum_{k=1}^n \sum_{\mathbf{m}_{k-1} \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{m}_k \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_k})}} \ln(A_{\mathbf{m}_{k-1}, \mathbf{m}_k}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \\
 & \times \left[\sum_{\mathbf{m}_0 \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_0})}} \dots \sum_{\mathbf{m}_{k-2} \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_{k-2}})}} \sum_{\mathbf{m}_{k+1} \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_{k+1}})}} \dots \sum_{\mathbf{m}_n \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_n})}} \mathbb{P}(\mathbf{m}_{t_0} = \mathbf{m}_0, \dots, \mathbf{m}_{t_n} = \mathbf{m}_n | \mathbf{Z}; \vartheta) \right] \\
 = & \sum_{k=1}^n \sum_{\mathbf{m}_{k-1} \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{m}_k \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_k})}} \ln(A_{\mathbf{m}_{k-1}, \mathbf{m}_k}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{m}_k, \mathbf{m}_{t_{k-1}} = \mathbf{m}_{k-1} | \mathbf{Z}; \vartheta)
 \end{aligned}$$

par la formule des probabilités totales appliquée aux variables $\mathbf{m}_{t_0}, \dots, \mathbf{m}_{t_{k-2}}, \mathbf{m}_{t_{k+1}}, \dots, \mathbf{m}_{t_n}$.

De la même manière, on obtient pour le dernier terme de la somme (F.1) :

$$\begin{aligned}
 & \sum_{\mathbf{m} \in \bigotimes_{i=0}^n \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_i})}} \mathbb{P}(\mathbf{M}_{t_0:t_n} = \mathbf{m} | \mathbf{Z}; \vartheta) \sum_{k=1}^n \ln(B(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{m}_k; \tilde{\vartheta})) \\
 & = \sum_{k=1}^n \sum_{\mathbf{m}_k \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{m}_k | \mathbf{Z}; \vartheta) \ln(B(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{m}_k; \tilde{\vartheta})).
 \end{aligned}$$

Ainsi,

$$\begin{aligned}
 Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= C(\vartheta, \mathbf{Z}, \Delta) + \sum_{k=1}^n \sum_{\mathbf{m}_k \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{m}_k | \mathbf{Z}; \vartheta) \ln \left(B \left(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{m}_k; \tilde{\vartheta} \right) \right) \\
 &\quad + \sum_{k=1}^n \sum_{\mathbf{m}_{k-1} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{m}_k \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} \ln(A_{\mathbf{m}_{k-1}, \mathbf{m}_k}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{m}_k, \mathbf{m}_{t_{k-1}} = \mathbf{m}_{k-1} | \mathbf{Z}; \vartheta) \\
 &= C(\vartheta, \mathbf{Z}, \Delta) + \sum_{k=1}^n \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}; \vartheta) \ln \left(B \left(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta} \right) \right) \\
 &\quad + \sum_{k=1}^n \sum_{\mathbf{u} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} \ln(A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v}, \mathbf{m}_{t_{k-1}} = \mathbf{u} | \mathbf{Z}; \vartheta).
 \end{aligned}$$

□

Dans la proposition suivante, remarquons que les termes pour lesquels $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \notin \mathcal{I}^\Lambda$ sont négligeables.

Proposition 2. *Si la condition (C1) est satisfaite, on a :*

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})(1 + o_{\mathbb{P}}(1)) + C(\vartheta, \mathbf{Z}, \Delta),$$

$\Delta \rightarrow 0$

avec :

$$\begin{aligned}
 \tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(B \left(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta} \right) \right) \\
 &\quad + \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda} \sum_{\mathbf{u} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})).
 \end{aligned}$$

Démonstration. Par la proposition 1, on a :

$$\begin{aligned}
 Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda} \sum_{\mathbf{u} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v}, \mathbf{m}_{t_{k-1}} = \mathbf{u} | \mathbf{Z}; \vartheta) \ln(A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \\
 &\quad + \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}; \vartheta) \ln \left(B \left(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta} \right) \right) + C(\vartheta, \mathbf{Z}, \Delta) + R(\Delta, \mathbf{Z}, \vartheta),
 \end{aligned}$$

où :

$$\begin{aligned}
 R(\Delta, \mathbf{Z}, \vartheta) &= \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \notin \mathcal{I}^\Lambda} \sum_{\mathbf{u} \in \mathbb{M}^{(\mathbf{z}_{t_{k-1}})} \mathbf{v} \in \mathbb{M}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v}, \mathbf{m}_{t_{k-1}} = \mathbf{u} | \mathbf{Z}; \vartheta) \ln(A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \\
 &+ \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \notin \mathcal{I}^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}; \vartheta) \ln(B(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta})).
 \end{aligned}$$

Montrons que $R(\Delta, \mathbf{Z}, \vartheta)$ est négligeable devant $Q(\tilde{\vartheta}, \vartheta, \mathbf{Z})$.

$$\begin{aligned}
 \mathbb{P}\left(\frac{|R(\Delta, \mathbf{Z}, \vartheta)|}{|Q(\tilde{\vartheta}, \vartheta, \mathbf{Z})|} > \epsilon\right) &\leq \mathbb{P}\left(\frac{|R(\Delta, \mathbf{Z}, \vartheta)|}{|Q(\tilde{\vartheta}, \vartheta, \mathbf{Z})|} > 0\right) \\
 &\leq \mathbb{P}(|R(\Delta, \mathbf{Z}, \vartheta)| > 0) \\
 &\leq \mathbb{P}\left(\bigcup_{k=1}^n \{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \notin \mathcal{I}^\Lambda\}\right) \\
 &\leq \mathbb{P}\left(\bigcup_{k=1}^n \{\text{plus de 2 sauts entre } t_k \text{ et } t_{k+1}\}\right) \\
 &\leq \sum_{k=1}^n \mathbb{P}(\{\text{plus de 2 sauts entre } t_k \text{ et } t_{k+1}\}).
 \end{aligned}$$

Si on note N_Δ le nombre de sauts qu'il y a eu sur l'intervalle $[0, \Delta]$ et α^* tel que pour tout \mathbf{x} , $\alpha(\mathbf{x}) \leq \alpha^*$. Les autres notations sont les mêmes que dans le chapitre 3.

On a :

$$\begin{aligned}
 &\sum_{k=1}^n \mathbb{P}(\{\text{plus de 2 sauts entre } t_k \text{ et } t_{k+1}\}) \\
 &= \sum_{k=1}^n \mathbb{E}(\mathbb{P}_{\mathbf{X}_{t_k}}(N_\Delta \geq 2)) \\
 &= \sum_{k=1}^n \mathbb{E}(\mathbb{E}_{\mathbf{X}_{t_k}}(\mathbb{1}_{\{N_\Delta \geq 2\}})) \\
 &= \sum_{k=1}^n \mathbb{E}(\mathbb{E}_{\mathbf{X}_{t_k}}(\mathbb{1}_{\{T_1 < \Delta\}} \mathbb{1}_{\{T_2 < \Delta - T_1\}})) \\
 &= \sum_{k=1}^n \mathbb{E}(\mathbb{E}_{\mathbf{X}_{t_k}}(\mathbb{E}_{\mathbf{X}_{t_k}}(\mathbb{1}_{\{T_1 < \Delta\}} \mathbb{1}_{\{T_2 - T_1 < \Delta - T_1\}} | T_1, \mathbf{X}_{T_1}, \mathbf{Y}^{(1)}))) \\
 &= \sum_{k=1}^n \mathbb{E}(\mathbb{E}_{\mathbf{X}_{t_k}}(\mathbb{1}_{\{T_1 < \Delta\}} \int_0^{\Delta - T_1} \alpha(\mathbf{Y}_s^{(1)}) e^{-\int_0^s \alpha(\mathbf{Y}_u^{(1)}) du} ds)) \\
 &\leq \sum_{k=1}^n \mathbb{E}(\mathbb{E}_{\mathbf{X}_{t_k}}(\mathbb{1}_{\{T_1 < \Delta\}} \Delta \alpha^*)),
 \end{aligned}$$

et :

$$\begin{aligned}
 \mathbb{E}_{\mathbf{X}_{t_k}} \left(\mathbf{1}_{\{T_1 < \Delta\}} \right) &= \mathbb{E}_{\mathbf{X}_{t_k}} \left(\mathbb{E} \left(\mathbf{1}_{\{T_1 < \Delta\}} \mid \mathbf{Y}^{(0)} \right) \right) \\
 &= \mathbb{E}_{\mathbf{X}_{t_k}} \left(\int_0^\Delta \alpha(\mathbf{Y}_s^{(0)}) e^{-\int_0^s \alpha(\mathbf{Y}_u^{(0)}) du} ds \right) \\
 &\leq \mathbb{E}_{\mathbf{X}_{t_k}} (\Delta \alpha^*) \\
 &\leq \Delta \alpha^*,
 \end{aligned}$$

donc :

$$\sum_{k=1}^n \mathbb{P}(\{\text{plus de 2 sauts entre } t_k \text{ et } t_{k+1}\}) \leq n \Delta^2 (\alpha^*)^2.$$

Or l'intervalle $[0, t]$ est découpé en n intervalles de taille Δ donc $\Delta \times n = t$ donc :

$$\sum_{k=1}^n \mathbb{P}(\{\text{plus de 2 sauts entre } t_k \text{ et } t_{k+1}\}) \leq \Delta^2 (\alpha^*)^2 \frac{t}{\Delta} \xrightarrow{\Delta \rightarrow 0} 0.$$

□

F.3.2 Forme de Q dans le cas où ϑ n'intervient que dans le move et la mutation

Dans cette section, nous allons simplifier l'expression de Q donnée dans la proposition 2 dans le cas où le paramètre d'intérêt ϑ n'intervient pas dans d'autres fonctions caractéristiques du processus que l'intensité de mutation τ , le noyau de mutation k_τ et la vraisemblance du move L .

Proposition 3. *Si les conditions (C1) et (C2) sont satisfaites, alors :*

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) \left(1 + o_{\mathbb{P}} \underset{\Delta \rightarrow 0}{(1)} \right) + \tilde{C}(\vartheta, \mathbf{Z}, \Delta),$$

où $C(\vartheta, \mathbf{Z}, \Delta)$ ne dépend pas de $\tilde{\vartheta}$ et où :

$$\begin{aligned}
 \tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda} \sum_{\mathbf{u} \in \mathbb{M}_\emptyset^{(\mathbf{z}_{t_{k-1}})} \mathbf{v} \in \mathbb{M}_\emptyset^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \\
 &+ \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_1^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_\emptyset^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L \left(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta} \right) \right) \\
 &+ \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_2^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_\emptyset^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L \left(\mathbf{z}_{t_k}^{(\mathbf{z}_{t_{k-1}})} | \mathbf{z}_{t_{k-1}}, \mathbf{v}^{(\mathbf{z}_{t_{k-1}})}; \tilde{\vartheta} \right) \right) \\
 &+ \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_3^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_\emptyset^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L \left(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}^{(\mathbf{z}_{t_k})}, \mathbf{v}; \tilde{\vartheta} \right) \right).
 \end{aligned}$$

Démonstration. On va simplifier le premier terme de la fonction \tilde{Q} de la proposition 2, à savoir :

$$\sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_\emptyset^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(B \left(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta} \right) \right).$$

Comme dans cette somme $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda$, on va chercher à calculer $B(\mathbf{z} | \mathbf{y}, \mathbf{v}; \tilde{\vartheta})$ pour $(\mathbf{y}, \mathbf{z}) \in \mathcal{I}$ et $\mathbf{v} \in \mathbb{M}_\emptyset^{(\mathbf{z})}$. Comme $\mathcal{I}^\Lambda = \mathcal{I}_1^\Lambda \cup \mathcal{I}_2^\Lambda \cup \mathcal{I}_3^\Lambda$, par la formule des probabilités totales, on a :

$$\begin{aligned}
 B(\mathbf{z} | \mathbf{y}, \mathbf{v}; \tilde{\vartheta}) &= \mathbb{P}(\mathbf{z}_{t_k} = \mathbf{z} | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}; \tilde{\vartheta}) \\
 &= \mathbb{P}(\mathbf{z}_{t_k} = \mathbf{z} | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}, (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_1^\Lambda; \tilde{\vartheta}) \mathbb{P}((\mathbf{y}, \mathbf{z}) \in \mathcal{I}_1^\Lambda | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}; \tilde{\vartheta}) \\
 &\quad + \mathbb{P}(\mathbf{z}_{t_k} = \mathbf{z} | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}, (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_2^\Lambda; \tilde{\vartheta}) \mathbb{P}((\mathbf{y}, \mathbf{z}) \in \mathcal{I}_2^\Lambda | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}; \tilde{\vartheta}) \\
 &\quad + \mathbb{P}(\mathbf{z}_{t_k} = \mathbf{z} | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}, (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_3^\Lambda; \tilde{\vartheta}) \mathbb{P}((\mathbf{y}, \mathbf{z}) \in \mathcal{I}_3^\Lambda | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}; \tilde{\vartheta}).
 \end{aligned}$$

Comme $\mathbf{v} \in \mathbb{M}_\emptyset^{(\mathbf{z})}$, on a pour $j \in \{1, 2, 3\}$:

$$\mathbb{P}((\mathbf{y}, \mathbf{z}) \in \mathcal{I}_j^\Lambda | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}; \tilde{\vartheta}) = \mathbb{1}_{(\mathbf{y}, \mathbf{v}) \in \mathcal{I}_j^\Lambda \times \mathbb{M}},$$

donc ce terme ne dépend pas de $\tilde{\vartheta}$.

De plus, dans le terme :

$$\begin{aligned}
 &\mathbb{P}(\mathbf{z}_{t_k} = \mathbf{z} | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}, (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_2^\Lambda; \tilde{\vartheta}), \\
 &\quad (\text{resp. } \mathbb{P}(\mathbf{z}_{t_k} = \mathbf{z} | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}, (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_3^\Lambda; \tilde{\vartheta}),)
 \end{aligned}$$

la marque $\mathbf{m}_{t_k} = \mathbf{v}$ fait partie du conditionnement donc le calcul de cette probabilité ne fait intervenir qu'une partie du noyau de naissance (resp. du noyau de mort) (dans lequel $\tilde{\vartheta}$ n'intervient pas) et la vraisemblance L du mouvement pour toutes les particules présentes à la fois dans la configuration \mathbf{y} et dans la configuration \mathbf{z} , donc il existe une constante $\tilde{C}_2(\mathbf{z}, \mathbf{y}, \mathbf{v})$ (resp. $\tilde{C}_3(\mathbf{z}, \mathbf{y}, \mathbf{v})$) qui ne dépend pas de $\tilde{\vartheta}$ telle que :

$$\begin{aligned} \mathbb{P}(\mathbf{z}_{t_k} = \mathbf{z} | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}, (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_2^\Lambda; \tilde{\vartheta}) &= \tilde{C}_2(\mathbf{z}, \mathbf{y}, \mathbf{v}) L(\mathbf{z}^{(\mathbf{y})} | \mathbf{y}, \mathbf{v}^{(\mathbf{y})}; \tilde{\vartheta}). \\ (\text{resp. } \mathbb{P}(\mathbf{z}_{t_k} = \mathbf{z} | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}, (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_3^\Lambda; \tilde{\vartheta}) &= \tilde{C}_3(\mathbf{z}, \mathbf{y}, \mathbf{v}) L(\mathbf{z} | \mathbf{y}^{(\mathbf{z})}, \mathbf{v}; \tilde{\vartheta}). \end{aligned}$$

De même, comme la nouvelle marque \mathbf{m}_{t_k} est déjà donnée dans le conditionnement, on a : $\mathbb{P}(\mathbf{z}_{t_k} = \mathbf{z} | \mathbf{z}_{t_{k-1}} = \mathbf{y}, \mathbf{m}_{t_k} = \mathbf{v}, (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_1^\Lambda; \tilde{\vartheta}) = L(\mathbf{z} | \mathbf{y}, \mathbf{v}; \tilde{\vartheta})$.

Finalement, en notant $C_1(\mathbf{z}, \mathbf{y}, \mathbf{v})$, $C_2(\mathbf{z}, \mathbf{y}, \mathbf{v})$, $C_3(\mathbf{z}, \mathbf{y}, \mathbf{v})$ des constantes qui ne dépendent pas de $\tilde{\vartheta}$, on a pour tout $(\mathbf{y}, \mathbf{z}) \in \mathcal{I}^\Lambda$:

$$B(\mathbf{z} | \mathbf{y}, \mathbf{v}; \tilde{\vartheta}) = \begin{cases} C_1(\mathbf{z}, \mathbf{y}, \mathbf{v}) L(\mathbf{z} | \mathbf{y}, \mathbf{v}; \tilde{\vartheta}) & \text{si } (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_1^\Lambda, \\ C_2(\mathbf{z}, \mathbf{y}, \mathbf{v}) L(\mathbf{z}^{(\mathbf{y})} | \mathbf{y}, \mathbf{v}^{(\mathbf{y})}; \tilde{\vartheta}) & \text{si } (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_2^\Lambda, \\ C_3(\mathbf{z}, \mathbf{y}, \mathbf{v}) L(\mathbf{z} | \mathbf{y}^{(\mathbf{z})}, \mathbf{v}; \tilde{\vartheta}) & \text{si } (\mathbf{y}, \mathbf{z}) \in \mathcal{I}_3^\Lambda. \end{cases}$$

Ainsi :

$$\begin{aligned} & \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(B(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta}) \right) \\ = & \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_1^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} p(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(C_1(\mathbf{z}_{t_k}, \mathbf{z}_{t_{k-1}}, \mathbf{v}) L(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta}) \right) \\ + & \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_2^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} p(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(C_2(\mathbf{z}_{t_k}, \mathbf{z}_{t_{k-1}}, \mathbf{v}) L(\mathbf{z}_{t_k}^{(\mathbf{z}_{t_{k-1}})} | \mathbf{z}_{t_{k-1}}, \mathbf{v}^{(\mathbf{u})}; \tilde{\vartheta}) \right) \\ + & \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_3^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} p(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(C_3(\mathbf{z}_{t_k}, \mathbf{z}_{t_{k-1}}, \mathbf{v}) L(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}^{(\mathbf{z}_{t_k})}, \mathbf{v}; \tilde{\vartheta}) \right). \end{aligned}$$

En développant, on obtient :

$$\begin{aligned}
 & \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} p(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(B \left(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta} \right) \right) = \tilde{C}(\mathbf{Z}, \vartheta, \Delta) \\
 & + \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_1^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} p(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L \left(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta} \right) \right) \\
 & + \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_2^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} p(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L \left(\mathbf{z}_{t_k}^{(\mathbf{z}_{t_{k-1}})} | \mathbf{z}_{t_{k-1}}, \mathbf{v}^{(\mathbf{z}_{t_{k-1}})}; \tilde{\vartheta} \right) \right) \\
 & + \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_3^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} p(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L \left(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta} \right) \right), \quad (\text{F.2})
 \end{aligned}$$

où $\tilde{C}(\mathbf{Z}, \vartheta, \Delta)$ est une constante en $\tilde{\vartheta}$ qui vaut :

$$\begin{aligned}
 \tilde{C}(\mathbf{Z}, \vartheta, \Delta) &= \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_1^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} p(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(C_1(\mathbf{z}_{t_k}, \mathbf{z}_{t_{k-1}}, \mathbf{v}) \right) \\
 &+ \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_2^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} p(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(C_2(\mathbf{z}_{t_k}, \mathbf{z}_{t_{k-1}}, \mathbf{v}) \right) \\
 &+ \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_3^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\mathcal{O}}^{(\mathbf{z}_{t_k})}} p(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(C_3(\mathbf{z}_{t_k}, \mathbf{z}_{t_{k-1}}, \mathbf{v}) \right).
 \end{aligned}$$

On a donc :

$$\tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \tilde{C}(\vartheta, \mathbf{Z}, \Delta) + \tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}).$$

Par la proposition 2, on a donc :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = C(\vartheta, \mathbf{Z}, \Delta) + \tilde{C}(\vartheta, \mathbf{Z}, \Delta) + \tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) + R(\Delta, \mathbf{Z}, \vartheta),$$

où $R(\Delta, \mathbf{Z}, \vartheta)$ est négligeable devant $\tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})$.

Par le même raisonnement que celui de la preuve de la proposition 2, on peut montrer que $R(\Delta, \mathbf{Z}, \vartheta)$ est négligeable devant $\tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})$, donc :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \tilde{C}(\vartheta, \mathbf{Z}, \Delta) + \tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) + R(\Delta, \mathbf{Z}, \vartheta),$$

où $\tilde{C}(\vartheta, \mathbf{Z}, \Delta) = C(\vartheta, \mathbf{Z}, \Delta) + \tilde{C}(\vartheta, \mathbf{Z}, \Delta)$. □

F.3.3 Forme de Q dans le cas où les mouvements et les changements de régime sont indépendants

Dans cette section, nous allons simplifier l'expression de Q donnée dans la proposition 3 dans le cas où les mouvements des trajectoires et leurs changements de régime sont indépendants au sens de la définition 1 page 169.

Proposition 4. *Si les conditions (C1), (C2) et (C3) sont satisfaites, alors la fonction Q s'écrit :*

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})(1 + o_{\mathbb{P}}(1)) + \hat{C}(\vartheta, \mathbf{Z}, \Delta),$$

avec :

$$\begin{aligned} \hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= \sum_{\substack{k=1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^n \sum_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \sum_{v \in \mathbb{M}} \mathbb{P}(m_{t_k, j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(b^j \left(z_{t_k, j} | z_{t_{k-1}, j}, v; \tilde{\vartheta} \right) \right) \\ &+ \sum_{\substack{k=1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^n \sum_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \sum_{u \in \mathbb{M}} \sum_{v \in \mathbb{M}} \mathbb{P}(m_{t_{k-1}, j} = u, m_{t_k, j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(a_{u,v}(z_{t_{k-1}, j}; \tilde{\vartheta})), \end{aligned}$$

où $\hat{C}(\vartheta, \mathbf{Z}, \Delta)$ est une constante en $\tilde{\vartheta}$.

De plus, on a aussi :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \bar{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})(1 + o_{\mathbb{P}}(1)) + \hat{C}(\vartheta, \mathbf{Z}, \Delta),$$

avec :

$$\begin{aligned} \bar{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \sum_{u=1}^M \sum_{v=1}^M \mathbb{P}(m_{t_{k-1}, j} = u, m_{t_k, j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(a_{u,v}(z_{t_{k-1}, j}; \tilde{\vartheta})) \\ &+ \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \sum_{v=1}^M \mathbb{P}(m_{t_k, j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(b^j \left(z_{t_k, j} | z_{t_{k-1}, j}, v; \tilde{\vartheta} \right) \right), \end{aligned}$$

où pour chaque trajectoire j on définit :

- $j_b = \min\{i, j \in \mathcal{P}(\mathbf{z}_{t_i})\}$ l'instant de naissance de la trajectoire j
- $j_d = \max\{i, j \in \mathcal{P}(\mathbf{z}_{t_i})\}$ l'instant de mort de la trajectoire j .

Démonstration. On fixe $k \in \llbracket 1, n \rrbracket$. Si $(\mathbf{u}, \mathbf{v}) \in \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_{k-1}})} \times \mathbb{M}_{\varnothing}^{(\mathbf{z}_{t_k})}$ avec $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda$

alors $(\mathbf{u}, \mathbf{v}) \in \mathcal{I}^{\mathbb{M}}$. Comme les changements de régime sont indépendants, par la définition 1, on a pour tout $(\mathbf{u}, \mathbf{v}) \in \mathcal{I}^{\mathbb{M}}$, pour tout $\mathbf{z} \in (\Lambda_{\emptyset})^s$ tel que $(\mathbf{z}, \mathbf{u}) \in \mathcal{I}_1^{\Lambda \times \mathbb{M}}$:

$$A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) \underset{\Delta \rightarrow 0}{\sim} C(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) \prod_{j=1}^s a_{u_j, v_j}(z_j; \tilde{\vartheta}).$$

Donc pour tout $k \in \llbracket 1, n \rrbracket$ tel que $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^{\Lambda}$, comme $(\mathbf{u}, \mathbf{v}) \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})} \times \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}$, on a :

$$A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta}) = \left(C(\mathbf{z}_{t_{k-1}}, \mathbf{u}, \mathbf{v}, \Delta) \prod_{j=1}^s a_{u_j, v_j}(z_{t_{k-1}, j}; \tilde{\vartheta}) \right) (1 + \underset{\Delta \rightarrow 0}{o}(1)),$$

donc :

$$\ln(A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) = \ln(C(\mathbf{z}_{t_{k-1}}, \mathbf{u}, \mathbf{v}, \Delta)) + \sum_{j=1}^s \ln(a_{u_j, v_j}(z_{t_{k-1}, j}; \tilde{\vartheta})) + \underset{\Delta \rightarrow 0}{o}(1).$$

Ainsi

$$\begin{aligned} & \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \\ &= \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \\ & \quad \times \left(\ln(C(\mathbf{z}_{t_{k-1}}, \mathbf{u}, \mathbf{v}, \Delta)) + \sum_{j=1}^s \ln(a_{u_j, v_j}(z_{t_{k-1}, j}; \tilde{\vartheta})) \right) \\ &+ \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \times \underset{\Delta \rightarrow 0}{o}(1). \end{aligned}$$

Donc :

$$\begin{aligned} & \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^{\Lambda}} \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}_{t_{k-1}}; \tilde{\vartheta})) \\ &= \tilde{C}(\vartheta, \mathbf{Z}, \Delta) + \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^{\Lambda}} \sum_{\substack{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})} \\ \mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})} \\ \mathbf{v} \neq \mathbf{u}}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \times \underset{\Delta \rightarrow 0}{o}(1) \\ &+ \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^{\Lambda}} \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \sum_{j=1}^s \ln(a_{u_j, v_j}(z_{t_{k-1}, j}; \tilde{\vartheta})), \end{aligned}$$

avec :

$$\tilde{C}(\vartheta, \mathbf{Z}, \Delta) = \sum_{\substack{k=1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^n \sum_{\mathbf{u} \in \mathbb{M}_{\circ}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\circ}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \\ \times \ln \left(C(\mathbf{z}_{t_{k-1}}, \mathbf{u}, \mathbf{v}, \Delta) \right).$$

Occupons nous d'abord du dernier terme de l'égalité précédente.

- Traitons d'abord le cas où $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_1^\Lambda$:

Si on note $i_1, \dots, i_{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})}$ les indices des coordonnées non égales à \circ dans le vecteur $\mathbf{z}_{t_{k-1}}$ (qui sont les mêmes que dans \mathbf{z}_{t_k} car $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_1^\Lambda$), alors on peut remplacer les sommes $\sum_{\mathbf{u} \in \mathbb{M}_{\circ}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\circ}^{(\mathbf{z}_{t_k})}}$ par les sommes $\sum_{u_{i_1} \in \mathbb{M}} \sum_{v_{i_1} \in \mathbb{M}} \dots \sum_{u_{i_{\#\mathcal{P}(\mathbf{z}_{t_k})}} \in \mathbb{M}} \sum_{v_{i_{\#\mathcal{P}(\mathbf{z}_{t_k})}} \in \mathbb{M}}$

et la somme $\sum_{j=1}^s$ par $\sum_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_k})}$.

On peut aussi remplacer

$$\mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) = \mathbb{P} \left(\bigcap_{j=1}^s (m_{t_{k-1},j} = u_j, m_{t_k,j} = v_j) | \mathbf{Z}_{t_0:t_n}, \vartheta \right)$$

par

$$\mathbb{P} \left(\bigcap_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_k})} (m_{t_{k-1},i_j} = u_{i_j}, m_{t_k,i_j} = v_{i_j}) | \mathbf{Z}_{t_0:t_n}; \vartheta \right)$$

car si on connaît $\mathbf{Z}_{t_0:t_n}$, et que $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_1^\Lambda$, alors on sait que $m_{t_{k-1},i_j} = m_{t_k,i_j} = \circ$ pour tout $j \notin \mathcal{P}(\mathbf{z}_{t_{k-1}}) = \mathcal{P}(\mathbf{z}_{t_k})$.

Ainsi :

$$\sum_{\mathbf{u} \in \mathbb{M}_{\circ}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\circ}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \sum_{j=1}^s \ln \left(a_{u_j, v_j}(z_{t_{k-1},j}; \tilde{\vartheta}) \right) \\ = \sum_{u_{i_1} \in \mathbb{M}} \sum_{v_{i_1} \in \mathbb{M}} \dots \sum_{u_{i_{\#\mathcal{P}(\mathbf{z}_{t_k})}} \in \mathbb{M}} \sum_{v_{i_{\#\mathcal{P}(\mathbf{z}_{t_k})}} \in \mathbb{M}} \mathbb{P} \left(\bigcap_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_k})} (m_{t_{k-1},i_j} = u_{i_j}, m_{t_k,i_j} = v_{i_j}) | \mathbf{Z}_{t_0:t_n}, \vartheta \right) \\ \times \sum_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_k})} \ln \left(a_{u_{i_j}, v_{i_j}}(z_{t_{k-1},i_j}; \tilde{\vartheta}) \right).$$

Faisons le calcul pour le facteur dans lequel $j = 1$, cela sera exactement pareil pour les autres valeurs de $j \in \llbracket 2, \#\mathcal{P}(\mathbf{z}_{t_k}) \rrbracket$.

$$\begin{aligned}
 & \sum_{u_{i_1} \in \mathbb{M}} \sum_{v_{i_1} \in \mathbb{M}} \cdots \sum_{u_{i_{\#\mathcal{P}(\mathbf{z}_{t_k})}} \in \mathbb{M}} \sum_{v_{i_{\#\mathcal{P}(\mathbf{z}_{t_k})}} \in \mathbb{M}} \mathbb{P} \left(\bigcap_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_k})} (m_{t_{k-1}, i_j} = u_j, m_{t_k, i_j} = v_j) \mid \mathbf{Z}_{t_0:t_n}, \vartheta \right) \\
 & \quad \times \ln \left(a_{u_{i_1}, v_{i_1}}(z_{t_{k-1}, i_1}; \tilde{\vartheta}) \right) \\
 & = \sum_{u_{i_1} \in \mathbb{M}} \sum_{v_{i_1} \in \mathbb{M}} \ln \left(a_{u_{i_1}, v_{i_1}}(z_{t_{k-1}, i_1}; \tilde{\vartheta}) \right) \\
 & \quad \times \sum_{u_{i_2} \in \mathbb{M}} \sum_{v_{i_2} \in \mathbb{M}} \cdots \sum_{u_{i_{\#\mathcal{P}(\mathbf{z}_{t_k})}} \in \mathbb{M}} \sum_{v_{i_{\#\mathcal{P}(\mathbf{z}_{t_k})}} \in \mathbb{M}} \mathbb{P} \left(\bigcap_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_k})} (m_{t_{k-1}, i_j} = u_j, m_{t_k, i_j} = v_j) \mid \mathbf{Z}_{t_0:t_n}, \vartheta \right) \\
 & = \sum_{u_{i_1} \in \mathbb{M}} \sum_{v_{i_1} \in \mathbb{M}} \ln \left(a_{u_{i_1}, v_{i_1}}(z_{t_{k-1}, i_1}; \tilde{\vartheta}) \right) \times \mathbb{P} \left(m_{t_{k-1}, i_1} = u_{i_1}, m_{t_k, i_1} = v_{i_1} \mid \mathbf{Z}_{t_0:t_n}, \vartheta \right)
 \end{aligned}$$

par la formule des probabilités totales appliquée aux variables aléatoires :

$$(m_{t_{k-1}, i_2}, m_{t_k, i_2}, \dots, m_{t_{k-1}, i_{\#\mathcal{P}(\mathbf{z}_{t_k})}}, m_{t_k, i_{\#\mathcal{P}(\mathbf{z}_{t_k})}}).$$

Ainsi :

$$\begin{aligned}
 & \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} \mid \mathbf{Z}_{t_0:t_n}, \vartheta) \sum_{j=1}^s \ln \left(a_{u_j, v_j}(z_{t_{k-1}, j}; \tilde{\vartheta}) \right) \\
 & = \sum_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_k})} \sum_{u_{i_j} \in \mathbb{M}} \sum_{v_{i_j} \in \mathbb{M}} \mathbb{P} \left(m_{t_{k-1}, i_j} = u_{i_j}, m_{t_k, i_j} = v_{i_j} \mid \mathbf{Z}_{t_0:t_n}, \vartheta \right) \ln \left(a_{u_{i_j}, v_{i_j}}(z_{t_{k-1}, i_j}; \tilde{\vartheta}) \right),
 \end{aligned}$$

donc :

$$\begin{aligned}
 & \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} \mid \mathbf{Z}_{t_0:t_n}, \vartheta) \sum_{j=1}^s \ln \left(a_{u_j, v_j}(z_{t_{k-1}, j}; \tilde{\vartheta}) \right) \\
 & = \sum_{j \in \mathcal{P}(\mathbf{z}_{t_{k-1}}) \cap \mathcal{P}(\mathbf{z}_{t_k})} \sum_{u \in \mathbb{M}} \sum_{v \in \mathbb{M}} \mathbb{P} \left(m_{t_{k-1}, j} = u, m_{t_k, j} = v \mid \mathbf{Z}_{t_0:t_n}, \vartheta \right) \ln \left(a_{u, v}(z_{t_{k-1}, j}; \tilde{\vartheta}) \right).
 \end{aligned} \tag{F.3}$$

• Si $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_2^\Lambda$, on a $\mathcal{P}(\mathbf{z}_{t_{k-1}}) \subset \mathcal{P}(\mathbf{z}_{t_k})$, et on procède exactement de la même manière en considérant les indices $i_1, \dots, i_{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})}$ des coordonnées non égales à \emptyset dans $\mathbf{z}_{t_{k-1}}$ et i_0 la coordonnée non égale à \emptyset dans \mathbf{z}_{t_k} mais égale à \emptyset dans $\mathbf{z}_{t_{k-1}}$.

Cela nous donne donc :

$$\begin{aligned}
 & \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \sum_{j=1}^s \ln \left(a_{u_j, v_j}(z_{t_{k-1}, j}; \tilde{\vartheta}) \right) \\
 &= \sum_{u_{i_1} \in \mathbb{M}} \sum_{v_{i_1} \in \mathbb{M}} \dots \sum_{u_{i_{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})}} \in \mathbb{M}} \sum_{v_{i_{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})}} \in \mathbb{M}} \sum_{v_{i_0} \in \mathbb{M}} \mathbb{P} \left(\bigcap_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})} (m_{t_{k-1}, i_j} = u_{i_j}, m_{t_k, i_j} = v_{i_j}) \cap \dots \right. \\
 & \quad \left. \dots \cap (m_{t_k, i_0} = v_{i_0}) | \mathbf{Z}_{t_0:t_n}, \vartheta \right) \sum_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})} \ln \left(a_{u_{i_j}, v_{i_j}}(z_{t_{k-1}, i_j}; \tilde{\vartheta}) \right).
 \end{aligned}$$

Pour $j = 1$, on a :

$$\begin{aligned}
 & \sum_{u_{i_1} \in \mathbb{M}} \sum_{v_{i_1} \in \mathbb{M}} \dots \sum_{u_{i_{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})}} \in \mathbb{M}} \sum_{v_{i_{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})}} \in \mathbb{M}} \sum_{v_{i_0} \in \mathbb{M}} \mathbb{P} \left(\bigcap_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})} (m_{t_{k-1}, i_j} = u_{i_j}, m_{t_k, i_j} = v_{i_j}) | \mathbf{Z}_{t_0:t_n}, \vartheta \right) \\
 & \quad \times \ln \left(a_{u_{i_1}, v_{i_1}}(z_{t_{k-1}, i_1}; \tilde{\vartheta}) \right) \\
 &= \sum_{u_{i_1} \in \mathbb{M}} \sum_{v_{i_1} \in \mathbb{M}} \ln \left(a_{u_{i_1}, v_{i_1}}(z_{t_{k-1}, i_1}; \tilde{\vartheta}) \right) \\
 & \times \sum_{u_{i_2} \in \mathbb{M}} \sum_{v_{i_2} \in \mathbb{M}} \dots \sum_{u_{i_{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})}} \in \mathbb{M}} \sum_{v_{i_{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})}} \in \mathbb{M}} \sum_{v_{i_0} \in \mathbb{M}} \mathbb{P} \left(\bigcap_{j=1}^{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})} (m_{t_{k-1}, i_j} = u_{i_j}, m_{t_k, i_j} = v_{i_j}) | \mathbf{Z}_{t_0:t_n}, \vartheta \right) \\
 &= \sum_{u_{i_1} \in \mathbb{M}} \sum_{v_{i_1} \in \mathbb{M}} \ln \left(a_{u_{i_1}, v_{i_1}}(z_{t_{k-1}, i_1}; \tilde{\vartheta}) \right) \times \mathbb{P} \left(m_{t_{k-1}, i_1} = u_{i_1}, m_{t_k, i_1} = v_{i_1} | \mathbf{Z}_{t_0:t_n}, \vartheta \right)
 \end{aligned}$$

par la formule des probabilités totales appliquée aux variables aléatoires :

$$(m_{t_{k-1}, i_2}, m_{t_k, i_2}, \dots, m_{t_{k-1}, i_{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})}}, m_{t_k, i_{\#\mathcal{P}(\mathbf{z}_{t_{k-1}})}}, m_{t_k, i_0}).$$

En faisant la même chose pour les autres valeurs de $j \in \llbracket 2, \#\mathcal{P}(\mathbf{z}_{t_k}) \rrbracket$, on retrouve (F.3).

• Si $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_3^{\Lambda}$, on a $\mathcal{P}(\mathbf{z}_{t_k}) \subset \mathcal{P}(\mathbf{z}_{t_{k-1}})$, et on réitère le même procédé en considérant les indices $i_1, \dots, i_{\#\mathcal{P}(\mathbf{z}_{t_k})}$ des coordonnées non égales à \emptyset dans \mathbf{z}_{t_k} et i_0 la coordonnée non égale à \emptyset dans $\mathbf{z}_{t_{k-1}}$ mais égale à \emptyset dans \mathbf{z}_{t_k} .

Exactement avec les mêmes techniques, mais en utilisant la définition d'indépen-

dance des mouvements qui nous dit que :

$$L(\mathbf{z}|\mathbf{y}, \mathbf{v}; \tilde{\vartheta}) = \prod_{\substack{j=1 \\ j \in \mathcal{P}(\mathbf{z}) \cap \mathcal{P}(\mathbf{y})}}^s b^j(z_j|y_j, v_j; \tilde{\vartheta}),$$

on obtient que :

- si $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_1^\Lambda$,

$$\begin{aligned} & \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta}) \right) \\ &= \sum_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \sum_{v \in \mathbb{M}} \mathbb{P}(m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(b^j(z_{t_k,j} | z_{t_{k-1},j}, v; \tilde{\vartheta}) \right), \end{aligned}$$

- si $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_2^\Lambda$,

$$\begin{aligned} & \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L(\mathbf{z}_{t_k}^{(\mathbf{z}_{t_{k-1}})} | \mathbf{z}_{t_{k-1}}, \mathbf{v}^{(\mathbf{z}_{t_{k-1}})}; \tilde{\vartheta}) \right) \\ &= \sum_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \sum_{v \in \mathbb{M}} \mathbb{P}(m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(b^j(z_{t_k,j} | z_{t_{k-1},j}, v; \tilde{\vartheta}) \right), \end{aligned}$$

- si $(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_3^\Lambda$,

$$\begin{aligned} & \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}^{(\mathbf{z}_{t_k})}, \mathbf{v}; \tilde{\vartheta}) \right) \\ &= \sum_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \sum_{v \in \mathbb{M}} \mathbb{P}(m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(b^j(z_{t_k,j} | z_{t_{k-1},j}, v; \tilde{\vartheta}) \right). \end{aligned}$$

En regroupant ces trois résultats, on obtient :

$$\begin{aligned} & \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_1^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}, \mathbf{v}; \tilde{\vartheta}) \right) \\ &+ \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_2^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L(\mathbf{z}_{t_k}^{(\mathbf{z}_{t_{k-1}})} | \mathbf{z}_{t_{k-1}}, \mathbf{v}^{(\mathbf{z}_{t_{k-1}})}; \tilde{\vartheta}) \right) \\ &+ \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}_3^\Lambda} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(L(\mathbf{z}_{t_k} | \mathbf{z}_{t_{k-1}}^{(\mathbf{z}_{t_k})}, \mathbf{v}; \tilde{\vartheta}) \right) \\ &= \sum_{k=1}^n \sum_{(\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda} \sum_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \sum_{v \in \mathbb{M}} \mathbb{P}(m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln \left(b^j(z_{t_k,j} | z_{t_{k-1},j}, v; \tilde{\vartheta}) \right). \end{aligned}$$

Ainsi, l'expression de \tilde{Q} donnée par la proposition 3 devient :

$$\begin{aligned} \tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= \tilde{C}(\vartheta, \mathbf{Z}, \Delta) + \hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) \\ &+ \sum_{\substack{k=1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^n \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\substack{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})} \\ \mathbf{v} \neq \mathbf{u}}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \times \underset{\Delta \rightarrow 0}{o}(1). \end{aligned} \quad (\text{F.4})$$

Démontrons que le dernier terme de cette égalité (F.4), i.e. la quantité :

$$\tilde{R}(\Delta, \mathbf{Z}, \vartheta) := \sum_{\substack{k=1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^n \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) \times f(\Delta),$$

où $f(\Delta) = \underset{\Delta \rightarrow 0}{o}(1)$, est négligeable devant $\hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = Q_a(\tilde{\vartheta}, \vartheta, \mathbf{Z}) + Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z})$ avec :

$$\begin{aligned} Q_a(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &:= \\ &\sum_{\substack{k=1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^n \sum_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \sum_{u \in \mathbb{M}} \sum_{v \in \mathbb{M}} \mathbb{P}(m_{t_{k-1},j} = u, m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(a_{u,v}(z_{t_{k-1},j}; \tilde{\vartheta})) \end{aligned}$$

et :

$$Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z}) := \sum_{\substack{k=1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^n \sum_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \sum_{v \in \mathbb{M}} p(m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(b^j(z_{t_k,j} | z_{t_{k-1},j}, v; \tilde{\vartheta})).$$

Tout d'abord, montrons que $\tilde{R}(\Delta, \mathbf{Z}, \vartheta)$ est négligeable devant $Q_a(\tilde{\vartheta}, \vartheta, \mathbf{Z})$. Comme les coefficients $a_{u,v}(z_{t_{k-1},j}) \in [0, 1]$, $Q_a(\tilde{\vartheta}, \vartheta, \mathbf{Z}) \leq 0$ p.s., en notant :

$$\bar{R}(\Delta, \mathbf{Z}, \vartheta) = \sum_{\substack{k=1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^n \sum_{\mathbf{u} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_{k-1}})}} \sum_{\mathbf{v} \in \mathbb{M}_{\emptyset}^{(\mathbf{z}_{t_k})}} \mathbb{P}(\mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{m}_{t_k} = \mathbf{v} | \mathbf{Z}_{t_0:t_n}, \vartheta) |f(\Delta)|,$$

on a :

$$\begin{aligned} \mathbb{P}\left(\left|\frac{\tilde{R}(\Delta, \mathbf{Z}, \vartheta)}{Q_a(\tilde{\vartheta}, \vartheta, \mathbf{Z})}\right| > \epsilon\right) &\leq \mathbb{P}\left(\frac{\bar{R}(\Delta, \mathbf{Z}, \vartheta)}{-Q_a(\tilde{\vartheta}, \vartheta, \mathbf{Z})} > \epsilon\right) \\ &= \mathbb{P}\left(\bar{R}(\Delta, \mathbf{Z}, \vartheta) + \epsilon Q_a(\tilde{\vartheta}, \vartheta, \mathbf{Z}) > 0\right). \end{aligned}$$

Or comme les coefficients $a_{u,v}(z_{t_{k-1},j}; \tilde{\vartheta})$ sont de la forme $e^{-\tau\Delta}$ ou $(1 - e^{-\tau\Delta})p_{u,v}$ (proposition 9 du chapitre 5) alors on a $Q_a(\tilde{\vartheta}, \vartheta, \mathbf{Z}) \xrightarrow[\Delta \rightarrow 0]{\mathbb{P}} -\infty$.

De plus, $f(\Delta) \xrightarrow{\Delta \rightarrow 0} 0$, donc :

$$\mathbb{P} \left(\frac{|\tilde{R}(\Delta, \mathbf{Z}, \vartheta)|}{|Q_a(\tilde{\vartheta}, \vartheta, \mathbf{Z})|} > \epsilon \right) \xrightarrow{\Delta \rightarrow 0} 0.$$

Montrons maintenant que $\tilde{R}(\Delta, \mathbf{Z}, \vartheta)$ est négligeable devant $Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z})$. Comme $Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z})$ n'est pas positif p.s., on procède différemment :

$$\mathbb{P} \left(\frac{|\tilde{R}(\Delta, \mathbf{Z}, \vartheta)|}{|Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z})|} > \epsilon \right) \leq \mathbb{P} \left(\bar{R}(\Delta, \mathbf{Z}, \vartheta) - \epsilon |Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z})| > 0 \right).$$

Or :

$$\begin{aligned} & \mathbb{P} \left(\bar{R}(\Delta, \mathbf{Z}, \vartheta) - \epsilon |Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z})| > 0 \right) \\ &= \mathbb{P} \left(\bar{R}(\Delta, \mathbf{Z}, \vartheta) - \epsilon |Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z})| > 0, \bigcap_{k=1}^n \bigcap_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \bigcap_{v \in \mathbb{M}} \left\{ \ln \left(b^j(z_{t_k, j} | z_{t_{k-1}, j}, v; \tilde{\vartheta}) \right) < 0 \right\} \right) \\ &+ \mathbb{P} \left(\bar{R}(\Delta, \mathbf{Z}, \vartheta) - \epsilon |Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z})| > 0, \bigcup_{k=1}^n \bigcup_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \bigcup_{v \in \mathbb{M}} \left\{ \ln \left(b^j(z_{t_k, j} | z_{t_{k-1}, j}, v; \tilde{\vartheta}) \right) \geq 0 \right\} \right) \\ &\leq \mathbb{P} \left(\bigcup_{k=1}^n \bigcup_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \bigcup_{v \in \mathbb{M}} \left\{ \ln \left(b^j(z_{t_k, j} | z_{t_{k-1}, j}, v; \tilde{\vartheta}) \right) < 0 \right\} \right) \\ &\quad + \mathbb{P} \left(\bar{R}(\Delta, \mathbf{Z}, \vartheta) - \epsilon |Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z})| > 0 \right). \end{aligned}$$

Montrons que les coefficients $b^j(z_{t_k, j} | z_{t_{k-1}, j}, v; \tilde{\vartheta})$ définis (cf proposition 10 du chapitre 5) comme :

$$b^i(z_{t_k, j} | z_{t_{k-1}, j}, v; \vartheta) = \begin{cases} \frac{1}{2\pi\sigma_{i,1}^2\Delta} \exp \left(-\frac{\|z_{t_k, j} - z_{t_{k-1}, j}\|^2}{2\sigma_{i,1}^2\Delta} \right) & \text{si } v = 1, \\ \frac{1}{2\pi\sigma_{i,2}^2\Delta} \exp \left(-\frac{\left\| z_{t_k, j} - z_{t_{k-1}, j} - \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta \right\|^2}{2\sigma_{i,2}^2\Delta} \right) & \text{si } v = 2, \\ \frac{1}{2\pi\sigma_{i,3}^2v(\lambda_i)} \exp \left(-\frac{\|z_{t_k, j} - e^{-\lambda_i\Delta}z_{t_{k-1}, j} - \mu_i(1 - e^{-\lambda_i\Delta})\|^2}{2\sigma_{i,3}^2v(\lambda_i)} \right) & \text{si } v = 3. \end{cases}$$

où $v(\lambda_i) = \frac{1}{2\lambda_i} (1 - e^{-2\lambda_i\Delta})$, convergent presque sûrement vers $+\infty$ lorsque $\Delta \rightarrow 0$.

Si $v = 1$, l'expression de $b^i(z_{t_k,j}|z_{t_{k-1},j}, 1; \vartheta)$ a été obtenue dans la preuve de la proposition 10 du chapitre 5 en utilisant le fait que $(z_{t,i})_{t \geq 0}$ est solution de l'EDS

$$dX_t = \sigma_{i,1} dB_t.$$

En intégrant cette EDS entre t_{k-1} et t_k (avec $\Delta = |t_k - t_{k-1}|$), on a :

$$z_{t_k,i} - z_{t_{k-1},i} = \sigma_{i,1} B_\Delta.$$

Par la loi du logarithme itéré pour un mouvement brownien dans \mathbb{R}^2 , on a avec probabilité 1 :

$$\limsup_{\Delta \rightarrow 0} \frac{\|z_{t_k,j} - z_{t_{k-1},j}\|}{\sqrt{2\sigma_{i,1}^2 \Delta \ln(\ln(1/\Delta))}} = 1,$$

donc pour Δ suffisamment petit, on a :

$$\begin{aligned} - \left\| \frac{z_{t_k,j} - z_{t_{k-1},j}}{2\sigma_{i,1}^2 \Delta} \right\|^2 &= - \left\| \frac{z_{t_k,j} - z_{t_{k-1},j}}{\sqrt{2\sigma_{i,1}^2 \Delta \ln(\ln(1/\Delta))}} \right\|^2 \times \ln(\ln(1/\Delta)) \\ &\geq -2 \ln(\ln(1/\Delta)), \end{aligned}$$

donc pour Δ suffisamment petit, on a :

$$\frac{1}{2\pi\sigma_{i,1}^2 \Delta} \exp\left(-\frac{\|z_{t_k,j} - z_{t_{k-1},j}\|^2}{2\sigma_{i,1}^2 \Delta}\right) \geq \frac{1}{2\pi\sigma_{i,1}^2 \Delta \ln(1/\Delta)^2},$$

donc :

$$b^i(z_{t_k,j}|z_{t_{k-1},j}, 1; \vartheta) \xrightarrow[\Delta \rightarrow 0]{\mathbb{P}} +\infty.$$

On fait de même si $v = 2$ avec l'EDS

$$dX_t = \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} dt + \sigma_{i,2} dB_t,$$

qui s'intègre en :

$$z_{t_k,i} - z_{t_{k-1},i} = \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta + \sigma_{i,2} B_\Delta,$$

donc pour Δ suffisamment petit, on a :

$$\begin{aligned} - \left\| \frac{z_{t_k,j} - z_{t_{k-1},j} - \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta}{2\sigma_{i,2}^2 \Delta} \right\|^2 &= - \left\| \frac{z_{t_k,j} - z_{t_{k-1},j} - \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta}{\sqrt{2\sigma_{i,2}^2 \Delta \ln(\ln(1/\Delta))}} \right\|^2 \times \ln(\ln(1/\Delta)) \\ &\geq -2 \ln(\ln(1/\Delta)), \end{aligned}$$

donc :

$$b^i(z_{t_k,j} | z_{t_{k-1},j}, \mathbf{2}; \vartheta) \xrightarrow[\Delta \rightarrow 0]{\mathbb{P}} +\infty.$$

Si $v = \mathbf{3}$, on a :

$$z_{t_k,i} - z_{t_{k-1},i} e^{-\lambda_i \Delta} - \mu_i (1 - e^{-\lambda_i \Delta}) = \frac{\sigma_{i,3}}{\sqrt{2\lambda_i}} e^{-\lambda_i \Delta} B_{e^{2\lambda_i \Delta} - 1}.$$

Par la loi du logarithme itéré pour un mouvement brownien dans \mathbb{R}^2 , on a avec probabilité 1 :

$$\limsup_{\Delta \rightarrow 0} \frac{\|B_{e^{2\lambda_i \Delta} - 1}\|}{\sqrt{2(e^{2\lambda_i \Delta} - 1) \ln(\ln((e^{2\lambda_i \Delta} - 1)^{-1}))}} = 1,$$

donc pour Δ suffisamment petit, on a :

$$\begin{aligned} - \frac{\|z_{t_k,j} - e^{-\lambda_i \Delta} z_{t_{k-1},j} - \mu_i (1 - e^{-\lambda_i \Delta})\|^2}{2\sigma_{i,3}^2 v(\lambda_i)} &= - \frac{\ln(\ln((e^{2\lambda_i \Delta} - 1)^{-1})) \|B_{e^{2\lambda_i \Delta} - 1}\|^2}{|2(e^{2\lambda_i \Delta} - 1) \ln(\ln((e^{2\lambda_i \Delta} - 1)^{-1}))|} \\ &\geq -2 \ln(\ln((e^{2\lambda_i \Delta} - 1)^{-1})), \end{aligned}$$

donc pour Δ suffisamment petit, on a :

$$\begin{aligned} b^i(z_{t_k,j} | z_{t_{k-1},j}, \mathbf{3}; \vartheta) &= \frac{1}{2\pi\sigma_{i,3}^2 v(\lambda_i)} \exp\left(-\frac{\|z_{t_k,j} - e^{-\lambda_i \Delta} z_{t_{k-1},j} - \mu_i (1 - e^{-\lambda_i \Delta})\|^2}{2\sigma_{i,3}^2 v(\lambda_i)}\right) \\ &\geq \frac{1}{2\pi\sigma_{i,3}^2 v(\lambda_i) \ln((e^{2\lambda_i \Delta} - 1)^{-1})^2} \\ &= \frac{1}{\pi\sigma_{i,3}^2 \ln(2\lambda_i \Delta)^2 \left(1 + \underset{\Delta \rightarrow 0}{o}\left(\frac{2}{\ln(2\lambda_i \Delta)}\right)\right) 2\lambda_i \Delta \left(1 + \underset{\Delta \rightarrow 0}{o}(1)\right)} \\ &\xrightarrow[\Delta \rightarrow 0]{} +\infty. \end{aligned}$$

Finalement, pour tout $v \in \mathbb{M}$,

$$b^i(z_{t_k,j}|z_{t_{k-1},j}, v; \vartheta) \xrightarrow[\Delta \rightarrow 0]{\mathbb{P}} +\infty.$$

Donc :

$$\mathbb{P} \left(\bigcup_{k=1}^n \bigcup_{j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})} \bigcup_{v \in \mathbb{M}} \{ \ln (b^j(z_{t_k,j}|z_{t_{k-1},j}, v; \tilde{\vartheta})) < 0 \} \right) \xrightarrow[\Delta \rightarrow 0]{} 0,$$

et :

$$Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z}) \xrightarrow[\Delta \rightarrow 0]{\mathbb{P}} +\infty.$$

De plus, $f(\Delta) \xrightarrow[\Delta \rightarrow 0]{} 0$, donc :

$$\mathbb{P} \left(\bar{R}(\Delta, \mathbf{Z}, \vartheta) - \epsilon Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z}) > 0 \right) \xrightarrow[\Delta \rightarrow 0]{} 0,$$

donc :

$$\mathbb{P} \left(\left| \frac{\tilde{R}(\Delta, \mathbf{Z}, \vartheta)}{Q_b(\tilde{\vartheta}, \vartheta, \mathbf{Z})} \right| > \epsilon \right) \xrightarrow[\Delta \rightarrow 0]{} 0.$$

Par (F.4), on a donc :

$$\tilde{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \tilde{C}(\vartheta, \mathbf{Z}, \Delta) + \hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) + \tilde{R}(\Delta, \mathbf{Z}, \vartheta),$$

où $\tilde{R}(\Delta, \mathbf{Z}, \vartheta)$ est négligeable devant $\hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})$.

Par la proposition 3, on a donc :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = C(\vartheta, \mathbf{Z}, \Delta) + \tilde{C}(\vartheta, \mathbf{Z}, \Delta) + \hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) + \tilde{R}(\Delta, \mathbf{Z}, \vartheta) + R(\Delta, \mathbf{Z}, \vartheta),$$

où $R(\Delta, \mathbf{Z}, \vartheta)$ est négligeable devant $\hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})$.

Par le même raisonnement que celui de la preuve de la proposition 2, on peut montrer que $R(\Delta, \mathbf{Z}, \vartheta)$ est négligeable devant $\hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})$, donc :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \hat{C}(\vartheta, \mathbf{Z}, \Delta) + \hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) + \hat{R}(\Delta, \mathbf{Z}, \vartheta),$$

avec $\hat{C}(\vartheta, \mathbf{Z}, \Delta) = \tilde{C}(\vartheta, \mathbf{Z}, \Delta) + C(\vartheta, \mathbf{Z}, \Delta)$, et $\hat{R}(\Delta, \mathbf{Z}, \vartheta) = R(\Delta, \mathbf{Z}, \vartheta) + \tilde{R}(\Delta, \mathbf{Z}, \vartheta)$,

qui est négligeable devant $\hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})$, donc :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \hat{C}(\vartheta, \mathbf{Z}, \Delta) + \hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})(1 + o_{\mathbb{P}}(1)).$$

Prouvons maintenant que $Q(\tilde{\vartheta}, \vartheta, \mathbf{Z})$ se réécrit de la façon suivante :

$$\begin{aligned} \hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \sum_{u=1}^M \sum_{v=1}^M \mathbb{P}(m_{t_{k-1},j} = u, m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(a_{u,v}(z_{t_{k-1},j}; \tilde{\vartheta})) \\ &+ \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \sum_{v=1}^M \mathbb{P}(m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(b^j(z_{t_k,j} | z_{t_{k-1},j}, v; \tilde{\vartheta})). \end{aligned}$$

Soit $j \in \llbracket 1, s \rrbracket$ et $k \in \llbracket 1, n \rrbracket$, on a :

$$k \in \llbracket j_b + 1, j_d \rrbracket \iff j \in \mathcal{P}(\mathbf{z}_{t_{k-1}}) \cap \mathcal{P}(\mathbf{z}_{t_k}).$$

Ainsi, les deux ensembles ci-dessous sont égaux :

$$\{(k, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, s \rrbracket, (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda \text{ et } k \in \llbracket j_b + 1, j_d \rrbracket\}$$

et

$$\{(k, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, s \rrbracket, (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda \text{ et } j \in \mathcal{P}(\mathbf{z}_{t_k}) \cap \mathcal{P}(\mathbf{z}_{t_{k-1}})\}.$$

C'est pourquoi on peut remplacer dans l'expression donnée par la proposition 4 :

$$\sum_{\substack{k=1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^n \sum_{j \in \mathcal{P}(\mathbf{z}_{t_{k-1}}) \cap \mathcal{P}(\mathbf{z}_{t_k})} \text{ par } \sum_{j=1}^s \sum_{\substack{k=j_b+1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}}}^{j_d}$$

et on obtient :

$$\begin{aligned} \hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) &= \sum_{j=1}^s \sum_{\substack{k=j_b+1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^{j_d} \sum_{u=1}^M \sum_{v=1}^M \mathbb{P}(m_{t_{k-1},j} = u, m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(a_{u,v}(z_{t_{k-1},j}; \tilde{\vartheta})) \\ &+ \sum_{j=1}^s \sum_{\substack{k=j_b+1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}^\Lambda}}^{j_d} \sum_{v=1}^M \mathbb{P}(m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(b^j(z_{t_k,j} | z_{t_{k-1},j}, v; \tilde{\vartheta})). \end{aligned}$$

Maintenant, on remarque que :

$$\begin{aligned} & \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \sum_{u=1}^M \sum_{v=1}^M \mathbb{P}(m_{t_{k-1},j} = u, m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(a_{u,v}(z_{t_{k-1},j}; \tilde{\vartheta})) \\ &= \sum_{j=1}^s \sum_{\substack{k=j_b+1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}}}^{j_d} \sum_{u=1}^M \sum_{v=1}^M \mathbb{P}(m_{t_{k-1},j} = u, m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(a_{u,v}(z_{t_{k-1},j}; \tilde{\vartheta})) + R_1(\Delta, \mathbf{Z}, \vartheta), \end{aligned}$$

avec

$$R_1(\Delta, \mathbf{Z}, \vartheta) = \sum_{j=1}^s \sum_{\substack{k=j_b+1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \notin \mathcal{I}}}^{j_d} \sum_{u=1}^M \sum_{v=1}^M \mathbb{P}(m_{t_{k-1},j} = u, m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(a_{u,v}(z_{t_{k-1},j}; \tilde{\vartheta})).$$

De même, on remarque que :

$$\begin{aligned} & \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \sum_{v=1}^M \mathbb{P}(m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(b^j(z_{t_k,j} | z_{t_{k-1},j}, v; \tilde{\vartheta})) \\ &= \sum_{j=1}^s \sum_{\substack{k=j_b+1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \in \mathcal{I}}}^{j_d} \sum_{v=1}^M \mathbb{P}(m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(b^j(z_{t_k,j} | z_{t_{k-1},j}, v; \tilde{\vartheta})) + R_2(\Delta, \mathbf{Z}, \vartheta), \end{aligned}$$

avec

$$R_2(\Delta, \mathbf{Z}, \vartheta) = \sum_{j=1}^s \sum_{\substack{k=j_b+1 \\ (\mathbf{z}_{t_{k-1}}, \mathbf{z}_{t_k}) \notin \mathcal{I}}}^{j_d} \sum_{v=1}^M \mathbb{P}(m_{t_k,j} = v | \mathbf{Z}_{t_0:t_n}, \vartheta) \ln(b^j(z_{t_k,j} | z_{t_{k-1},j}, v; \tilde{\vartheta})).$$

Donc :

$$\hat{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \bar{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z}) + R_1(\Delta, \mathbf{Z}, \vartheta) + R_2(\Delta, \mathbf{Z}, \vartheta).$$

On peut montrer avec le même raisonnement que celui de la preuve de la proposition 2, que $R_1(\Delta, \mathbf{Z}, \vartheta)$, $R_2(\Delta, \mathbf{Z}, \vartheta)$ et $\hat{R}(\Delta, \mathbf{Z}, \vartheta)$ sont négligeables devant $\bar{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})$, donc :

$$Q(\tilde{\vartheta}, \vartheta, \mathbf{Z}) = \hat{C}(\vartheta, \mathbf{Z}, \Delta) + \bar{Q}(\tilde{\vartheta}, \vartheta, \mathbf{Z})(1 + o_{\mathbb{P}}(1)).$$

□

F.4 Indépendance des mouvements des trajectoires et des changements de régime dans l'exemple de la partie 5.3 du chapitre 5

Montrons que dans l'exemple de la section 5.3 du chapitre 5, les trois conditions **(C1)**, **(C2)** et **(C3)** sont vérifiées.

Dans cet exemple, les intensités de mort δ et de naissance β sont bornées. De plus, il existe $n^* \in \mathbb{N}$ tel que :

$$\tau(\mathbf{x}) = \tau \times n(\mathbf{x}) \leq \tau \times n^*,$$

donc α est bornée donc **(C1)** est vérifiée.

Le paramètre d'intérêt

$$\vartheta = (\tau, p_b, p_r, p_v, \sigma_{1,1}^2, \rho_1, \eta_1, \sigma_{1,2}^2, \lambda_1, \mu_1, \sigma_{1,3}^2, \dots, \sigma_{s,1}^2, \rho_s, \eta_s, \sigma_{s,2}^2, \lambda_s, \mu_s, \sigma_{s,3}^2)$$

intervient par définition seulement dans les fonctions τ , k_τ et L , donc **(C2)** est vérifiée.

Il nous reste donc à montrer **(C3)**.

Ainsi, montrons que pour tout $(\mathbf{u}, \mathbf{v}) \in \mathcal{I}^{\mathbb{M}}$, pour tout $\mathbf{z} \in (\Lambda_\emptyset)^s$ tel que pour tout $j \in \llbracket 1, s \rrbracket$, $z_j = \emptyset$ si $u_j = \emptyset$, on a :

$$A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) \underset{\Delta \rightarrow 0}{\sim} C(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) \prod_{j \in \mathcal{P}(\mathbf{u}) \cap \mathcal{P}(\mathbf{v})}^s a_{u_j, v_j}(z_j; \tilde{\vartheta})$$

où $C(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta)$ est une constante qui ne dépend pas de $\tilde{\vartheta}$ et où $a_{u,v}(z; \tilde{\vartheta})$ est défini dans l'équation (5.1) (section 5.2.1 du chapitre 5)

On commence par calculer $\prod_{j \in \mathcal{P}(\mathbf{u}) \cap \mathcal{P}(\mathbf{v})}^s a_{u_j, v_j}(z_j; \tilde{\vartheta})$:

Par la proposition 9 du chapitre 5, $a_{u_j, v_j}(z_j; \tilde{\vartheta})$ ne dépend pas de z_j et est équivalent au coefficient (u_j, v_j) de la matrice

$$\begin{array}{c}
 1 \qquad \qquad \qquad 2 \qquad \qquad \qquad 3 \\
 \begin{array}{l}
 1 \\
 2 \\
 3
 \end{array}
 \left(\begin{array}{ccc}
 e^{-\tau\Delta} & p_1(1 - e^{-\tau\Delta}) & (1 - p_1)(1 - e^{-\tau\Delta}) \\
 p_2(1 - e^{-\tau\Delta}) & e^{-\tau\Delta} & (1 - p_2)(1 - e^{-\tau\Delta}) \\
 p_3(1 - e^{-\tau\Delta}) & (1 - p_3)(1 - e^{-\tau\Delta}) & e^{-\tau\Delta}
 \end{array} \right)
 \end{array}$$

donc :

$$\prod_{\substack{j=1 \\ j \in \mathcal{P}(\mathbf{z})}}^s a_{u_j, v_j}(z_j; \tilde{\vartheta}) \underset{\Delta \rightarrow 0}{\sim} e^{-\Delta\tau \#\{j \in \mathcal{P}(\mathbf{z}), u_i = v_i\}} (1 - e^{-\Delta\tau})^{\#\{j \in \mathcal{P}(\mathbf{z}), u_j \neq v_j\}} \prod_{\substack{j=1 \\ u_j \neq v_j}}^s p_{u_j, v_j}$$

donc :

$$\prod_{\substack{j=1 \\ j \in \mathcal{P}(\mathbf{z})}}^s a_{u_j, v_j}(z_j; \tilde{\vartheta}) \underset{\Delta \rightarrow 0}{\sim} \begin{cases} e^{-\Delta\tau(\#\mathcal{P}(\mathbf{z})-q)} (1 - e^{-\Delta\tau})^q \prod_{\substack{j=1 \\ u_j \neq v_j}}^s p_{u_j, v_j} & \text{si } \#\{j \in \mathcal{P}(\mathbf{z}), u_j \neq v_j\} = q \geq 1, \\ e^{-\Delta\tau \#\mathcal{P}(\mathbf{z})} & \text{si } \forall j \in \mathcal{P}(\mathbf{z}), u_j = v_j. \end{cases}$$

Ainsi :

$$\prod_{\substack{j=1 \\ j \in \mathcal{P}(\mathbf{z})}}^s a_{u_j, v_j}(z_j; \tilde{\vartheta}) \underset{\Delta \rightarrow 0}{\sim} \begin{cases} (\Delta\tau)^q \prod_{\substack{j=1 \\ u_j \neq v_j}}^s p_{u_j, v_j} & \text{si } \#\{j \in \mathcal{P}(\mathbf{z}), u_j \neq v_j\} = q \geq 1, \\ e^{-\Delta\tau \#\mathcal{P}(\mathbf{z})} & \text{si } \forall j \in \mathcal{P}(\mathbf{z}), u_j = v_j. \end{cases}$$

Maintenant, calculons $A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}; \tilde{\vartheta})$ pour $(\mathbf{u}, \mathbf{v}) \in \mathcal{I}^{\mathbb{M}}$:

Soit $(\mathbf{u}, \mathbf{v}) \in \mathcal{I}^{\mathbb{M}}$, alors $|\#\mathcal{P}(\mathbf{u}) - \#\mathcal{P}(\mathbf{v})| \leq 1$. On sépare donc les deux cas suivants :

- **Cas où $\#\mathcal{P}(\mathbf{u}) = \#\mathcal{P}(\mathbf{v})$:**

Comme $(\mathbf{u}, \mathbf{v}) \in \mathcal{I}^{\mathbb{M}}$ alors $\mathcal{P}(\mathbf{u}) = \mathcal{P}(\mathbf{v}) = \mathcal{P}(\mathbf{z})$. On va alors calculer :

$$A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) = \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta})$$

en séparant selon 2 cas :

- ✧ Premier cas : si $\#\{i \in \mathcal{P}(\mathbf{z}), u_i \neq v_i\} \geq 1$,
- ✧ Deuxième cas : si $\#\{i \in \mathcal{P}(\mathbf{z}), u_i \neq v_i\} = 0$.

Traisons ces deux cas :

✧ Premier cas : Si on note $\#\{i \in \mathcal{P}(\mathbf{z}), u_i \neq v_i\} = q \geq 1$, alors on sait qu'il y a eu au moins q mutations, donc au moins q sauts entre t_{k-1} et t_k , et on a donc :

$$\begin{aligned} A_{\mathbf{u},\mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) &= \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}) \\ &= \mathbb{P}\left(\mathbf{m}_{t_k} = \mathbf{v}, \left\{ \begin{array}{l} q \text{ sauts entre } t_{k-1} \text{ et } t_k \\ \text{qui sont des mutations} \end{array} \right\} \middle| \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}\right) \\ &\quad + \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v}, \{\text{plus de } q + 1 \text{ sauts entre } t_{k-1} \text{ et } t_k\} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}). \end{aligned}$$

Or :

$$\mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v}, \{\text{plus de } q + 1 \text{ sauts entre } t_{k-1} \text{ et } t_k\} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}) = \underset{\Delta \rightarrow 0}{o}(\Delta^q).$$

Et :

$$\begin{aligned} &\mathbb{P}\left(\mathbf{m}_{t_k} = \mathbf{v}, \left\{ \begin{array}{l} q \text{ sauts entre } t_{k-1} \text{ et } t_k \\ \text{qui sont des mutations} \end{array} \right\} \middle| \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}\right) \\ &= \mathbb{P}\left(\mathbf{m}_{t_k} = \mathbf{v}, \left\{ \begin{array}{l} q \text{ sauts entre } t_{k-1} \text{ et } t_k \\ \text{qui sont des mutations} \end{array} \right\}, \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}\right) \\ &\quad \times \mathbb{P}\left(\left\{ \begin{array}{l} q \text{ sauts entre } t_{k-1} \text{ et } t_k \\ \text{qui sont des mutations} \end{array} \right\} \middle| \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}\right). \end{aligned}$$

Par la forme du noyau de mutation donnée en section 5.3.1 du chapitre 5, on a :

$$\begin{aligned} &\mathbb{P}\left(\mathbf{m}_{t_k} = \mathbf{v}, \left\{ \begin{array}{l} q \text{ sauts entre } t_{k-1} \text{ et } t_k \\ \text{qui sont des mutations} \end{array} \right\}, \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}\right) \\ &= \frac{1}{(\#\mathcal{P}(\mathbf{z}))^q} \prod_{\substack{j=1 \\ u_j \neq v_j}}^s p_{u_j, v_j}. \end{aligned}$$

Comme les q sauts qui ont lieu entre t_{k-1} et t_k sont tous des mutations, le nombre de particules présentes ne change pas (et est égal à $\#\mathcal{P}(\mathbf{z})$) entre t_{k-1} et t_k . Ainsi, le nombre de mutations entre t_{k-1} et t_k suit une loi exponentielle de paramètre $\Delta\tau\#\mathcal{P}(\mathbf{z})$, on a :

$$\mathbb{P}\left(\left\{ \begin{array}{l} q \text{ sauts entre } t_{k-1} \text{ et } t_k \\ \text{qui sont des mutations} \end{array} \right\} \middle| \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}\right) = \frac{(\tau\Delta\#\mathcal{P}(\mathbf{z}))^q e^{-\Delta\tau\#\mathcal{P}(\mathbf{z})}}{q!},$$

donc finalement :

$$\begin{aligned}
A_{\mathbf{u},\mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) &= \frac{1}{(\#\mathcal{P}(\mathbf{z}))^q} \prod_{\substack{j=1 \\ u_j \neq v_j}}^s p_{u_j, v_j} \frac{(\tau \Delta \#\mathcal{P}(\mathbf{z}))^q e^{-\Delta \tau \#\mathcal{P}(\mathbf{z})}}{q!} + \underset{\Delta \rightarrow 0}{o}(\Delta^q) \\
&= \prod_{\substack{j=1 \\ u_j \neq v_j}}^s p_{u_j, v_j} \frac{(\tau \Delta)^q e^{-\Delta \tau \#\mathcal{P}(\mathbf{z})}}{q!} + \underset{\Delta \rightarrow 0}{o}(\Delta^q) \\
&\underset{\Delta \rightarrow 0}{\sim} C_1(\mathbf{z}, \mathbf{u}, \mathbf{v}) \prod_{\substack{j=1 \\ u_j \neq v_j}}^s p_{u_j, v_j} (\tau \Delta)^q.
\end{aligned}$$

◇ Deuxième cas : Si $\forall i \in \mathcal{P}(\mathbf{z}), u_i = v_i$, alors :

$$\begin{aligned}
A_{\mathbf{u},\mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) &= \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}) \\
&= \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}, \{\text{pas de saut entre } t_{k-1} \text{ et } t_k\}; \tilde{\vartheta}) \\
&\quad \times \mathbb{P}(\{\text{pas de saut entre } t_{k-1} \text{ et } t_k\} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}) \\
&\quad + \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v}, \{\text{plus de 2 sauts entre } t_{k-1} \text{ et } t_k\} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}).
\end{aligned}$$

On a :

$$\begin{aligned}
&\mathbb{P}(\{\text{pas de saut entre } t_{k-1} \text{ et } t_k\} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}) \\
&= \mathbb{E}_{(\mathbf{z}, \mathbf{u})} \left(\exp \left(- \int_0^\Delta \alpha((\mathbf{Y}_s^{(k-1)}, \mathbf{u}), \tilde{\vartheta}) ds \right) \right) \\
&= \mathbb{E}_{(\mathbf{z}, \mathbf{u})} \left(\exp \left(- \int_0^\Delta (\tau + \beta + \delta)((\mathbf{Y}_s^{(k-1)}, \mathbf{u}), \tilde{\vartheta}) ds \right) \right) \\
&= \mathbb{E}_{(\mathbf{z}, \mathbf{u})} \left(\exp \left(- \int_0^\Delta \tau n((\mathbf{Y}_s^{(k-1)}, \mathbf{u})) ds \right) \exp \left(- \int_0^\Delta (\beta + \delta)((\mathbf{Y}_s^{(k-1)}, \mathbf{u})) ds \right) \right) \\
&= \mathbb{E}_{(\mathbf{z}, \mathbf{u})} \left(\exp(-\Delta \tau n((\mathbf{z}, \mathbf{u}))) \exp \left(- \int_0^\Delta (\beta + \delta)((\mathbf{Y}_s^{(k-1)}, \mathbf{u})) ds \right) \right) \\
&= e^{-\tau \#\mathcal{P}(\mathbf{z}) \Delta} \mathbb{E}_{(\mathbf{z}, \mathbf{u})} \left(\exp \left(- \int_0^\Delta (\beta + \delta)((\mathbf{Y}_s^{(k-1)}, \mathbf{u})) ds \right) \right) \\
&= C_3(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) e^{-\tau \#\mathcal{P}(\mathbf{z}) \Delta},
\end{aligned}$$

avec :

$$C_3(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) = \mathbb{E}_{(\mathbf{z}, \mathbf{u})} \left(\exp \left(- \int_0^\Delta (\beta + \delta)((\mathbf{Y}_s^{(k-1)}, \mathbf{u})) ds \right) \right)$$

qui tend vers 1 quand $\Delta \rightarrow 0$ par convergence dominée (car α est bornée).

De plus, comme $\mathbf{u} = \mathbf{v}$, on a :

$$\mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}, \{\text{pas de saut entre } t_{k-1} \text{ et } t_k\}; \tilde{\vartheta}) = 1.$$

Ainsi :

$$\begin{aligned} A_{\mathbf{u},\mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) &= C_3(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) e^{-\tau \#\mathcal{P}(\mathbf{z})\Delta} + \underset{\Delta \rightarrow 0}{o}(\Delta) \\ &\underset{\Delta \rightarrow 0}{\sim} C_3(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) e^{-\tau \#\mathcal{P}(\mathbf{z})\Delta}. \end{aligned}$$

Ainsi :

$$A_{\mathbf{u},\mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) \underset{\Delta \rightarrow 0}{\sim} \begin{cases} C_1(\mathbf{z}, \mathbf{u}, \mathbf{v})(\tau\Delta)^q \prod_{\substack{j=1 \\ u_j \neq v_j}}^s p_{u_j, v_j} & \text{si } \#\{i \in \mathcal{P}(\mathbf{z}), u_i \neq v_i\} = q \geq 1, \\ C_3(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) e^{-\tau \#\mathcal{P}(\mathbf{z})\Delta} & \text{si } \forall i \in \mathcal{P}(\mathbf{z}), u_i = v_i. \end{cases}$$

Comme $C_3(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) \xrightarrow{\Delta \rightarrow 0} 1$, on a donc bien dans ce cas :

$$A_{\mathbf{u},\mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) \underset{\Delta \rightarrow 0}{\sim} C_4(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) \prod_{\substack{j=1 \\ j \in \mathcal{P}(\mathbf{u}) \cap \mathcal{P}(\mathbf{v})}}^s a_{u_j, v_j}(z_j; \tilde{\vartheta}).$$

- **Cas où $|\#\mathcal{P}(\mathbf{u}) - \#\mathcal{P}(\mathbf{v})| = 1$:**

On note i_0 tel que $\{i_0\} = \{i \in \llbracket 1, s \rrbracket, (u_{i_0} \neq \emptyset \text{ et } v_{i_0} = \emptyset) \text{ ou } (u_{i_0} = \emptyset \text{ et } v_{i_0} \neq \emptyset)\}$.

$$\begin{aligned} A_{\mathbf{u},\mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) &= \mathbb{P}(\mathbf{m}_{t_k} = \mathbf{v} | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}) \\ &= \mathbb{P}(m_{t_k, i_0} = v_{i_0} | \bigcap_{\substack{i=1 \\ i \neq i_0}}^s (m_{t_k, i} = v_i), \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}) \\ &\quad \times \mathbb{P}(\bigcap_{\substack{i=1 \\ i \neq i_0}}^s (m_{t_k, i} = v_i) | \mathbf{m}_{t_{k-1}} = \mathbf{u}, \mathbf{z}_{t_{k-1}} = \mathbf{z}; \tilde{\vartheta}) \\ &= C_5(\mathbf{z}, \mathbf{u}, \mathbf{v}) \Delta e^{-\tau((\mathbf{z}, \mathbf{u}); \tilde{\vartheta})\Delta} \times A_{\tilde{\mathbf{u}}, \tilde{\mathbf{v}}}(\mathbf{z}; \tilde{\vartheta}) \\ &\underset{\Delta \rightarrow 0}{\sim} C_5(\mathbf{z}, \mathbf{u}, \mathbf{v}) \Delta A_{\tilde{\mathbf{u}}, \tilde{\mathbf{v}}}(\mathbf{z}; \tilde{\vartheta}). \end{aligned}$$

où $\tilde{\mathbf{u}}$ et $\tilde{\mathbf{v}}$ sont les vecteurs \mathbf{u} et \mathbf{v} avec leur coordonnée i_0 égale à \emptyset , ie :

$$(\tilde{\mathbf{u}}, \tilde{\mathbf{v}}) = \begin{cases} (\mathbf{u}, \mathbf{v}^{(\mathbf{u})}) & \text{si } (v_{i_0} \neq \emptyset \text{ et } u_{i_0} = \emptyset), \\ (\mathbf{u}^{(\mathbf{v})}, \mathbf{v}) & \text{si } (u_{i_0} \neq \emptyset \text{ et } v_{i_0} = \emptyset). \end{cases}$$

Par le cas précédent, il existe une constante $C_6(\mathbf{z}, \tilde{\mathbf{u}}, \tilde{\mathbf{v}}, \Delta)$ en $\tilde{\vartheta}$ telle que :

$$A_{\tilde{\mathbf{u}}, \tilde{\mathbf{v}}}(\mathbf{z}; \tilde{\vartheta}) \underset{\Delta \rightarrow 0}{\sim} C_6(\mathbf{z}, \tilde{\mathbf{u}}, \tilde{\mathbf{v}}, \Delta) \prod_{\substack{j=1 \\ j \in \mathcal{P}(\tilde{\mathbf{u}}) \cap \mathcal{P}(\tilde{\mathbf{v}})}}^s a_{\tilde{u}_j, \tilde{v}_j}(z_j; \tilde{\vartheta}).$$

Donc :

$$\begin{aligned} A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) &= C_4(\mathbf{z}, \mathbf{u}, \mathbf{v}) \Delta C_6(\mathbf{z}, \tilde{\mathbf{u}}, \tilde{\mathbf{v}}, \Delta) \prod_{\substack{j=1 \\ j \in \mathcal{P}(\tilde{\mathbf{u}}) \cap \mathcal{P}(\tilde{\mathbf{v}})}}^s a_{\tilde{u}_j, \tilde{v}_j}(z_j; \tilde{\vartheta}) \\ &= C_7(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) \prod_{\substack{j=1 \\ j \in \mathcal{P}(\mathbf{u}) \cap \mathcal{P}(\mathbf{v})}}^s a_{u_j, v_j}(z_j; \tilde{\vartheta}). \end{aligned}$$

On a donc bien montré que pour tout $(\mathbf{u}, \mathbf{v}) \in \mathcal{I}^{\mathbb{M}}$, pour tout $\mathbf{z} \in (\Lambda_{\mathcal{O}})^s$ tel que pour tout $j \in \llbracket 1, s \rrbracket$, $z_j = \mathcal{O}$ si $u_j = \mathcal{O}$, on a :

$$A_{\mathbf{u}, \mathbf{v}}(\mathbf{z}; \tilde{\vartheta}) \underset{\Delta \rightarrow 0}{\sim} C_7(\mathbf{z}, \mathbf{u}, \mathbf{v}, \Delta) \prod_{\substack{j=1 \\ j \in \mathcal{P}(\mathbf{u}) \cap \mathcal{P}(\mathbf{v})}}^s a_{u_j, v_j}(z_j; \tilde{\vartheta}).$$

PREUVES DU CALCUL DE L'ARGUMENT DU MAXIMUM DE Q

Dans toute cette annexe, pour faciliter la rédaction, on notera, pour tout $i \in \llbracket 1, s \rrbracket$:

$$r_i := (\sigma_{i,1}^2, \rho_i, \eta_i, \sigma_{i,2}^2, \lambda_i, \mu_i, \sigma_{i,3}^2)$$

l'ensemble des paramètres qui se rapportent à la trajectoire i . On a $r_i \in R$ avec :

$$R := \mathbb{R}_+^* \times \mathbb{R}^* \times \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[\times \mathbb{R}_+^* \times \mathbb{R}_+ \times \Lambda \times \mathbb{R}_+^*.$$

G.1 Preuve de la proposition 11 du chapitre 5

Les paramètres $(\tau, p_b, p_r, p_v) \in T := \mathbb{R}^* \times]0, 1]^3$ interviennent dans Q seulement par la matrice a , ainsi :

$$\begin{aligned} & \arg \max_{(\tau, p_b, p_r, p_v) \in T} Q((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z}) \\ &= \arg \max_{(\tau, p_b, p_r, p_v) \in T} Q^1((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z}) \\ &= \arg \max_{(\tau, p_b, p_r, p_v) \in T} \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \sum_{u=1}^M \sum_{v=1}^M \xi_{k-1}^i(u, v; \vartheta) \ln(a_{u,v}(\tilde{\vartheta})). \end{aligned}$$

Or en distinguant les différents cas pour $u, v \in \llbracket 1, 3 \rrbracket$, on a :

$$\begin{aligned}
 \sum_{u=1}^M \sum_{v=1}^M \ln(a_{u,v}(\tau, p_b, p_r, p_v)) \xi_{k-1}^i(u, v; \vartheta) &= \ln(e^{-\tau\Delta}) \xi_{k-1}^i(1, 1; \vartheta) \\
 &+ \ln(p_b(1 - e^{-\tau\Delta})) \xi_{k-1}^i(1, 2; \vartheta) + \ln((1 - p_b)(1 - e^{-\tau\Delta})) \xi_{k-1}^i(1, 3; \vartheta) \\
 &+ \ln(p_r(1 - e^{-\tau\Delta})) \xi_{k-1}^i(2, 1; \vartheta) + \ln(e^{-\tau\Delta}) \xi_{k-1}^i(2, 2; \vartheta) \\
 &+ \ln((1 - p_r)(1 - e^{-\tau\Delta})) \xi_{k-1}^i(2, 3; \vartheta) + \ln(p_v(1 - e^{-\tau\Delta})) \xi_{k-1}^i(3, 1; \vartheta) \\
 &+ \ln((1 - p_v)(1 - e^{-\tau\Delta})) \xi_{k-1}^i(3, 2; \vartheta) + \ln(e^{-\tau\Delta}) \xi_{k-1}^i(3, 3; \vartheta).
 \end{aligned}$$

Donc on cherche l'argument du maximum de la fonction :

$$\begin{aligned}
 f : (\tau, p_b, p_r, p_v) \in T &\mapsto \ln(e^{-\tau\Delta})(C_{1,1} + C_{2,2} + C_{3,3}) + \ln(1 - e^{-\tau\Delta})C \\
 &+ \ln(p_b)C_{1,2} + \ln(1 - p_b)C_{1,3} + \ln(p_r)C_{2,1} \\
 &+ \ln(1 - p_r)C_{2,3} + \ln(p_v)C_{3,1} + \ln(1 - p_v)C_{3,2}
 \end{aligned}$$

où on note pour $(j, \ell) \in \llbracket 1, 3 \rrbracket^2$:

$$C_{j,\ell} = \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \xi_{k-1}^i(j, \ell; \vartheta) \quad \text{et} \quad C = C_{1,2} + C_{1,3} + C_{2,1} + C_{2,3} + C_{3,1} + C_{3,2}.$$

On calcule les dérivées de f par rapport aux variables τ, p_b, p_r et p_v :

$$\begin{aligned}
 \frac{\partial f}{\partial \tau}(\tau, p_b, p_r, p_v) &= -\Delta(C_{1,1} + C_{2,2} + C_{3,3}) + \frac{\Delta e^{-\tau\Delta}}{1 - e^{-\tau\Delta}} C, \\
 \frac{\partial f}{\partial p_b}(\tau, p_b, p_r, p_v) &= \frac{1}{p_b} C_{1,2} - \frac{1}{1 - p_b} C_{1,3}, \\
 \frac{\partial f}{\partial p_r}(\tau, p_b, p_r, p_v) &= \frac{1}{p_r} C_{2,1} - \frac{1}{1 - p_r} C_{2,3}, \\
 \frac{\partial f}{\partial p_v}(\tau, p_b, p_r, p_v) &= \frac{1}{p_v} C_{3,1} - \frac{1}{1 - p_v} C_{3,2}.
 \end{aligned}$$

En annulant ces dérivées, on obtient un seul point critique, que l'on note $(\bar{\tau}, \bar{p}_b, \bar{p}_r, \bar{p}_v)$:

$$\begin{aligned}
 \bar{\tau} &= \frac{1}{\Delta} \ln \left(1 + \frac{C}{C_{1,1} + C_{2,2} + C_{3,3}} \right), & \bar{p}_r &= \frac{C_{2,1}}{C_{2,1} + C_{2,3}}, \\
 \bar{p}_b &= \frac{C_{1,2}}{C_{1,2} + C_{1,3}}, & \bar{p}_v &= \frac{C_{3,1}}{C_{3,1} + C_{3,2}}.
 \end{aligned}$$

On vérifie maintenant que ce point critique est bien un maximum en calculant les dérivées secondes et en les évaluant en le point critique $(\bar{\tau}, \bar{p}_b, \bar{p}_r, \bar{p}_v)$, on obtient :

$$\begin{aligned}\frac{\partial^2 f}{\partial p_b^2}(\bar{\tau}, \bar{p}_b, \bar{p}_r, \bar{p}_v) &= -\frac{(C_{1,2} + C_{1,3})^3}{C_{1,2}C_{1,3}}, \\ \frac{\partial^2 f}{\partial p_r^2}(\bar{\tau}, \bar{p}_b, \bar{p}_r, \bar{p}_v) &= -\frac{(C_{2,1} + C_{2,3})^3}{C_{2,1}C_{2,3}}, \\ \frac{\partial^2 f}{\partial p_v^2}(\bar{\tau}, \bar{p}_b, \bar{p}_r, \bar{p}_v) &= -\frac{(C_{3,1} + C_{3,2})^3}{C_{3,1}C_{3,2}}, \\ \frac{\partial^2 f}{\partial \tau \partial p_i}(\bar{\tau}, \bar{p}_b, \bar{p}_r, \bar{p}_v) &= 0, \quad \forall i \in \{b, r, v\},\end{aligned}$$

et en se servant de :

$$\begin{aligned}\sum_{j=1}^3 \sum_{\ell=1}^3 C_{j,\ell} &= \sum_{j=1}^3 \sum_{\ell=1}^3 \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \xi_{k-1}^i(j, \ell; \vartheta) \\ &= \sum_{j=1}^3 \sum_{\ell=1}^3 \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} p(m_{t_{k-1},i} = j, m_{t_k,i} = \ell | \mathbf{Z}^i; \vartheta) \\ &= \sum_{i=1}^s (i_d - i_b),\end{aligned}$$

on obtient que :

$$\begin{aligned}\frac{\partial^2 f}{\partial \tau^2}(\bar{\tau}, \bar{p}_b, \bar{p}_r, \bar{p}_v) &= -\Delta^2 (C_{1,1} + C_{2,2} + C_{3,3}) \frac{C_{1,1} + C_{2,2} + C_{3,3} + C}{C} \\ &= -\Delta^2 \frac{\left(\sum_{i=1}^s (i_d - i_b) \right) (C_{1,2} + C_{2,2} + C_{3,3})}{C} =: a_{11}.\end{aligned}$$

La matrice hessienne de f en le point critique $(\bar{\tau}, \bar{p}_b, \bar{p}_r, \bar{p}_v)$ est donc :

$$\text{Hess}_{(\bar{\tau}, \bar{p}_b, \bar{p}_r, \bar{p}_v)} f = \begin{pmatrix} a_{11} & 0 & 0 & 0 \\ 0 & -\frac{(C_{1,2} + C_{1,3})^3}{C_{1,2}C_{1,3}} & 0 & 0 \\ 0 & 0 & -\frac{(C_{2,1} + C_{2,3})^3}{C_{2,1}C_{2,3}} & 0 \\ 0 & 0 & 0 & -\frac{(C_{3,1} + C_{3,2})^3}{C_{3,1}C_{3,2}} \end{pmatrix}.$$

Comme pour tout $(j, \ell) \in \llbracket 1, 3 \rrbracket^2$, $C_{j,\ell} > 0$, cette matrice est définie négative et le point critique $(\bar{\tau}, \bar{p}_b, \bar{p}_r, \bar{p}_v)$ est donc un maximum local. Comme f est \mathcal{C}^∞ sur l'ouvert T , cet unique point critique est donc un maximum global.

G.2 Preuve de la proposition 12 du chapitre 5

Les paramètres compris dans r_i interviennent dans Q seulement par la fonction b^i , ainsi :

$$\begin{aligned} \arg \max_{r_i \in \mathbb{R}^8} Q((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z}) \\ &= \arg \max_{r_i \in \mathbb{R}^8} Q^2((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z}) \\ &= \arg \max_{r_i \in \mathbb{R}^8} \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \sum_{v=1}^M \gamma_k^j(v, \vartheta) \ln \left(b^j \left(z_{t_k, j} | z_{t_{k-1}, j}, v, \tilde{\vartheta} \right) \right) \\ &= \arg \max_{r_i \in \mathbb{R}^8} \sum_{k=i_b+1}^{i_d} \sum_{v=1}^M \gamma_k^i(v, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, v, r_i \right) \right). \end{aligned}$$

Etant donné la forme de la fonction b^i (proposition 10 du chapitre 5), on a pour tout $k \in \llbracket i_b + 1, i_d \rrbracket$:

$$\begin{aligned} \sum_{v=1}^M \gamma_k^i(v, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, v, r_i \right) \right) &= \gamma_k^i(1, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, 1, r_i \right) \right) \\ &\quad + \gamma_k^i(2, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, 2, r_i \right) \right) + \gamma_k^i(3, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, 3, r_i \right) \right) \\ &= \gamma_k^i(1, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, 1, (\sigma_{i,1}^2) \right) \right) + \gamma_k^i(2, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, 2, (\rho_i, \eta_i, \sigma_{i,2}^2) \right) \right) \\ &\quad + \gamma_k^i(3, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, 3, (\lambda_i, \mu_i(0), \mu_i(1), \sigma_{i,3}^2) \right) \right). \end{aligned}$$

Ainsi, pour rechercher $\arg \max_{r_i \in \mathbb{R}^8} Q((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z})$, on peut rechercher séparément

$$\arg \max_{\sigma_{i,1}^2 \in \mathbb{R}_+^*} Q((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z}),$$

puis

$$\arg \max_{(\rho_i, \eta_i, \sigma_{i,2}^2) \in \mathbb{R}^* \times]-\frac{\pi}{2}, \frac{\pi}{2}[\times \mathbb{R}_+^*} Q((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z}),$$

puis

$$\arg \max_{(\lambda_i, (\mu_i(0), \mu_i(1)), \sigma_{i,3}^2) \in \mathbb{R}_+ \times \Lambda \times \mathbb{R}_+^*} Q((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z}).$$

On commence par $\arg \max_{\sigma_{i,1}^2 \in \mathbb{R}_+^*} Q((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z})$. Pour cela, on cherche le maximum de la fonction :

$$\sigma_{i,1}^2 \mapsto \sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, 1, (\sigma_{i,1}^2) \right) \right).$$

On a par la forme de $b^i(\cdot, 1, \cdot)$ donnée dans la proposition 10 du chapitre 5 :

$$\begin{aligned} & \sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, 1, (\sigma_{i,1}^2) \right) \right) \\ &= \sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta) \ln \left(\frac{1}{2\pi\sigma_{i,1}^2\Delta} \exp \left(-\frac{\|z_{t_k, i} - z_{t_{k-1}, i}\|^2}{2\sigma_{i,1}^2\Delta} \right) \right) \\ &= - \sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta) \ln(2\pi\sigma_{i,1}^2\Delta) - \sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta) \frac{\|z_{t_k, i} - z_{t_{k-1}, i}\|^2}{2\sigma_{i,1}^2\Delta}. \end{aligned}$$

La dérivée de cette fonction est :

$$\sigma_{i,1}^2 \mapsto \frac{1}{\sigma_{i,1}^2} \left(- \sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta) + \frac{1}{2\Delta\sigma_{i,1}^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2 \right).$$

Elle s'annule en :

$$\widehat{\sigma_{i,1}^2} = \frac{1}{2\Delta} \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta)},$$

et est positive sur $]0, \widehat{\sigma_{i,1}^2}]$ et négative sur $[\widehat{\sigma_{i,1}^2}, +\infty[$ donc la fonction présente un maximum global en $\widehat{\sigma_{i,1}^2}$.

Ensuite, pour trouver $\arg \max_{(\rho_i, \eta_i, \sigma_{i,2}^2) \in \mathbb{R}^* \times]-\frac{\pi}{2}, \frac{\pi}{2}[\times \mathbb{R}_+^*} Q((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z})$, on cherche le maximum de la fonction :

$$h : (\rho_i, \eta_i, \sigma_{i,2}^2) \mapsto \sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta) \ln \left(b^i \left(z_{t_k,i} | z_{t_{k-1},i}, 2, (\rho_i, \eta_i, \sigma_{i,2}^2) \right) \right).$$

On a par la forme de $b^i(\cdot, \cdot, 2, \cdot)$ donnée dans la proposition 10 du chapitre 5 :

$$\begin{aligned} & \sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta) \ln \left(b^i \left(z_{t_k,i} | z_{t_{k-1},i}, 2, (\rho_i, \eta_i, \sigma_{i,2}^2) \right) \right) \\ &= - \sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta) \ln \left(2\pi\sigma_{i,2}^2\Delta \right) - \sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta) \frac{\left\| z_{t_k,i} - z_{t_{k-1},i} - \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta \right\|^2}{2\sigma_{i,2}^2\Delta}. \end{aligned}$$

On calcule les dérivées par rapport aux variables ρ_i, η_i et $\sigma_{i,2}^2$:

$$\begin{aligned} \frac{\partial h}{\partial \sigma_{i,2}^2}(\rho_i, \eta_i, \sigma_{i,2}^2) &= -\frac{1}{\sigma_{i,2}^2} \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta) - \sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta) \left\| z_{t_k,i} - z_{t_{k-1},i} - \rho_i \begin{pmatrix} \cos(\eta_i) \\ \sin(\eta_i) \end{pmatrix} \Delta \right\|^2 \right), \\ \frac{\partial h}{\partial \eta_i}(\rho_i, \eta_i, \sigma_{i,2}^2) &= -\frac{\rho_i}{\sigma_{i,2}^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta) \left\langle z_{t_k} - z_{t_{k-1}}, \begin{pmatrix} \sin(\eta_i) \\ \cos(\eta_i) \end{pmatrix} \right\rangle, \\ \frac{\partial h}{\partial \rho_i}(\rho_i, \eta_i, \sigma_{i,2}^2) &= \frac{1}{\sigma_{i,2}^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta) \left(\left\langle z_{t_k} - z_{t_{k-1}}, \begin{pmatrix} \sin(\eta_i) \\ \cos(\eta_i) \end{pmatrix} \right\rangle - \rho_i \Delta \right). \end{aligned}$$

On annule ces dérivées et on obtient le point critique $(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2)$ avec :

$$\begin{aligned} \widehat{\eta}_i &= \arctan \left(\frac{\sum_{k=i_b+1}^{i_d} (z_{t_k,i}(1) - z_{t_{k-1},i}(1)) \gamma_k^i(2, \vartheta)}{\sum_{k=i_b+1}^{i_d} (z_{t_k,i}(0) - z_{t_{k-1},i}(0)) \gamma_k^i(2, \vartheta)} \right), \\ \widehat{\rho}_i &= \frac{1}{\Delta} \frac{\sum_{k=i_b+1}^{i_d} \left\langle \begin{pmatrix} \cos(\widehat{\eta}_i) \\ \sin(\widehat{\eta}_i) \end{pmatrix}, z_{t_k,i} - z_{t_{k-1},i} \right\rangle \gamma_k^i(2, \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta)}, \\ \widehat{\sigma}_{i,2}^2 &= \frac{1}{2\Delta} \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta) \left\| z_{t_k,i} - z_{t_{k-1},i} - \widehat{\rho}_i \begin{pmatrix} \cos(\widehat{\eta}_i) \\ \sin(\widehat{\eta}_i) \end{pmatrix} \Delta \right\|^2}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta)}. \end{aligned}$$

Afin de connaître la nature du point critique $(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2)$, on calcule la matrice hessienne de h au point $(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2)$. Pour cela, on calcule les dérivées secondes de h qu'on évalue au point $(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2)$. On a :

$$\frac{\partial^2 h}{\partial (\sigma_{i,2}^2)^2}(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2) = -\frac{1}{(\widehat{\sigma}_{i,2}^2)^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta),$$

$$\frac{\partial^2 h}{\partial \eta_i^2}(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2) = -\frac{\Delta \widehat{\rho}_i^2}{\widehat{\sigma}_{i,2}^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta),$$

$$\frac{\partial^2 h}{\partial \rho_i^2}(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2) = -\frac{1}{\widehat{\sigma}_{i,2}^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(2, \vartheta).$$

On a aussi :

$$\frac{\partial^2 h}{\partial \sigma_{i,2}^2 \partial \eta_i}(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2) = -\frac{1}{\widehat{\sigma}_{i,2}^2} \times \frac{\partial h}{\partial \eta_i}(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2) = 0,$$

$$\frac{\partial^2 h}{\partial \sigma_{i,2}^2 \partial \rho_i}(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2) = -\frac{1}{\widehat{\sigma}_{i,2}^2} \times \frac{\partial h}{\partial \rho_i}(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2) = 0,$$

$$\frac{\partial^2 h}{\partial \rho_i \partial \eta_i}(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2) = -\frac{1}{\widehat{\rho}_i} \times \frac{\partial h}{\partial \eta_i}(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2) = 0.$$

La matrice hessienne de h au point $(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2)$ est donc diagonale avec des coefficients diagonaux strictement négatifs. Elle est donc définie négative et le point critique $(\widehat{\rho}_i, \widehat{\eta}_i, \widehat{\sigma}_{i,2}^2)$ est donc un maximum local. Comme h est \mathcal{C}^∞ sur l'ouvert $\mathbb{R}^* \times]-\frac{\pi}{2}, \frac{\pi}{2}[\times \mathbb{R}_+^*$, cet unique point critique est donc un maximum global.

Enfin, pour trouver

$$\arg \max_{(\lambda_i, \mu_i(0), \mu_i(1), \sigma_{i,3}^2) \in \mathbb{R}_+ \times \Lambda \times \mathbb{R}_+^*} Q((\tau, p_b, p_r, p_v, r_1, \dots, r_s), \vartheta, \mathbf{Z}),$$

on cherche le maximum de la fonction :

$$(\lambda_i, \mu_i(0), \mu_i(1), \sigma_{i,3}^2) \mapsto \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, 3, (\lambda_i, \mu_i(0), \mu_i(1), \sigma_{i,3}^2) \right) \right).$$

On a par la forme de $b^i(\cdot|\cdot, 3, \cdot)$ donnée dans la proposition 10 du chapitre 5 :

$$\begin{aligned}
 & \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln \left(b^i \left(z_{t_k, i} | z_{t_{k-1}, i}, 3, (\lambda_i, \mu_i(0), \mu_i(1), \sigma_{i,3}^2) \right) \right) \\
 &= \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln \left(\frac{1}{2\pi\sigma_{i,3}^2 v(\lambda_i)} \exp \left(-\frac{\|z_{t_k, i} - z_{t_{k-1}, i} e^{-\lambda_i \Delta} - \mu_i(1 - e^{-\lambda_i \Delta})\|^2}{2\sigma_{i,3}^2 v(\lambda_i)} \right) \right) \\
 &= - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln \left(2\pi\sigma_{i,3}^2 v(\lambda_i) \right) - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{\|z_{t_k, i} - z_{t_{k-1}, i} e^{-\lambda_i \Delta} - \mu_i(1 - e^{-\lambda_i \Delta})\|^2}{2\sigma_{i,3}^2 v(\lambda_i)}.
 \end{aligned}$$

On fait le changement de variable :

$$\begin{cases} a_i = e^{-\lambda_i \Delta} \\ b_i = \sigma_{i,3}^2 v(\lambda_i) \\ \mu_i(0) = \mu_i(0) \\ \mu_i(1) = \mu_i(1) \end{cases}$$

et on obtient une fonction des variables $a_i, b_i, \mu_i(0), \mu_i(1)$, que l'on notera f . On a :

$$\begin{aligned}
 f(a_i, b_i, \mu_i(0), \mu_i(1)) &= - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln(2\pi b_i) \\
 &\quad - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{\|z_{t_k, i} - z_{t_{k-1}, i} a_i - \mu_i(1 - a_i)\|^2}{2b_i}.
 \end{aligned}$$

On calcule les dérivées par rapport aux variables $a_i, b_i, \mu_i(0), \mu_i(1)$ et on obtient :

$$\begin{aligned}
 \frac{\partial f}{\partial \mu_i(0)}(a_i, b_i, \mu_i(0), \mu_i(1)) &= - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{(z_{t_k, i}(0) - z_{t_{k-1}, i}(0) a_i - \mu_i(0)(1 - a_i))(a_i - 1)}{b_i}, \\
 \frac{\partial f}{\partial \mu_i(1)}(a_i, b_i, \mu_i(0), \mu_i(1)) &= - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{(z_{t_k, i}(1) - z_{t_{k-1}, i}(1) a_i - \mu_i(1)(1 - a_i))(a_i - 1)}{b_i}, \\
 \frac{\partial f}{\partial a_i}(a_i, b_i, \mu_i(0), \mu_i(1)) &= - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{\langle z_{t_k, i} - z_{t_{k-1}, i} a_i - \mu_i(1 - a_i), -z_{t_{k-1}, i} + \mu_i \rangle}{b_i}, \\
 \frac{\partial f}{\partial b_i}(a_i, b_i, \mu_i(0), \mu_i(1)) &= -\frac{1}{b_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) + \frac{1}{2b_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} a_i - \mu_i(1 - a_i)\|^2.
 \end{aligned} \tag{G.1}$$

On cherche les points critiques de cette fonction. On a :

$$\left\{ \begin{array}{l} \frac{\partial f}{\partial \mu_i(0)} = 0 \\ \frac{\partial f}{\partial \mu_i(1)} = 0 \end{array} \right. \iff \left\{ \begin{array}{l} a_i = 1 \\ \text{ou} \\ \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) (z_{t_k, i}(p) - z_{t_{k-1}, i}(p) a_i - \mu_i(p)(1 - a_i)) = 0, \forall p \in \{0, 1\} \end{array} \right.$$

(G.2)

$$\iff \left\{ \begin{array}{l} a_i = 1 \\ \text{ou} \\ \mu_i(p)(1 - a_i) = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) (z_{t_k, i}(p) - z_{t_{k-1}, i}(p) a_i)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)}, \forall p \in \{0, 1\} \end{array} \right.$$

(G.3)

Étudions maintenant ces deux cas possibles l'un après l'autre.

Étude du cas $a_i = 1$: On notera $(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)})$ le point critique.

Si $a_i = 1$, il reste à résoudre :

$$\left\{ \begin{array}{l} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \frac{\langle z_{t_k, i} - z_{t_{k-1}, i}, -z_{t_{k-1}, i} + \mu_i \rangle}{b_i} = 0 \\ -\frac{1}{b_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) + \frac{1}{2b_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2 = 0. \end{array} \right.$$

D'où :

$$\bar{b}_i = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2}{2 \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)}. \quad (\text{G.4})$$

Le vecteur $\overline{\mu}_i := \begin{pmatrix} \overline{\mu}_i(0) \\ \overline{\mu}_i(1) \end{pmatrix}$ est solution de l'équation :

$$\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \langle z_{t_k, i} - z_{t_{k-1}, i}, -z_{t_{k-1}, i} + \overline{\mu}_i \rangle = 0.$$

Afin de connaître la nature du point critique $(1, \overline{b}_i, \overline{\mu}_i(0), \overline{\mu}_i(1))$, étudions la matrice hessienne de la fonction f en ce point. Pour cela, on calcule les dérivées secondes grâce aux dérivées premières (G.1) de cette fonction f et on les évalue au point critique $(1, \overline{b}_i, \overline{\mu}_i(0), \overline{\mu}_i(1))$:

$$\frac{\partial^2 f}{\partial a_i^2}(1, \overline{b}_i, \overline{\mu}_i(0), \overline{\mu}_i(1)) = - \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \frac{\| -z_{t_{k-1}, i} + \overline{\mu}_i \|^2}{\overline{b}_i} =: a_{11}.$$

Et :

$$\frac{\partial^2 f}{\partial b_i^2}(1, \overline{b}_i, \overline{\mu}_i(0), \overline{\mu}_i(1)) = \frac{1}{\overline{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) - \frac{1}{\overline{b}_i^3} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2 =: a_{22}.$$

Ainsi en utilisant l'expression de \overline{b}_i donnée en (G.4), on a :

$$\begin{aligned} a_{22} &= \frac{1}{\overline{b}_i^2} \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) - \frac{1}{\overline{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2 \right) \\ &= \frac{1}{\overline{b}_i^2} \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) - 2 \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \right) \\ &= -\frac{1}{\overline{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta). \end{aligned}$$

On a aussi :

$$\frac{\partial^2 f}{\partial \mu_i(0)^2}(1, \overline{b}_i, \overline{\mu}_i(0), \overline{\mu}_i(1)) = 0.$$

De la même manière, on obtient que :

$$\frac{\partial^2 f}{\partial \mu_i(1)^2}(1, \overline{b}_i, \overline{\mu}_i(0), \overline{\mu}_i(1)) = 0.$$

Puis :

$$\frac{\partial^2 f}{\partial a_i \partial b_i}(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)}) = \frac{1}{b_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \langle z_{t_k, i} - z_{t_{k-1}, i}, -z_{t_{k-1}, i} + \bar{\mu}_i \rangle =: a_{12}.$$

Et :

$$\frac{\partial^2 f}{\partial a_i \partial \mu_i(0)}(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)}) = -\frac{1}{b_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) (z_{t_k, i}(0) - z_{t_{k-1}, i}(0)) =: a_{13}.$$

De la même manière on obtient :

$$\frac{\partial^2 f}{\partial a_i \partial \mu_i(1)}(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)}) = -\frac{1}{b_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) (z_{t_k, i}(1) - z_{t_{k-1}, i}(1)) =: a_{14}.$$

Aussi, on a :

$$\frac{\partial^2 f}{\partial b_i \partial \mu_i(0)}(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)}) = 0.$$

De la même manière on obtient :

$$\frac{\partial^2 f}{\partial b_i \partial \mu_i(1)}(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)}) = 0.$$

Enfin, on a :

$$\frac{\partial^2 f}{\partial \mu_i(0) \partial \mu_i(1)}(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)}) = 0.$$

La matrice hessienne de f en le point critique $(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)})$ est donc :

$$\text{Hess}_{(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)})} f = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{12} & a_{22} & 0 & 0 \\ a_{13} & 0 & 0 & 0 \\ a_{14} & 0 & 0 & 0 \end{pmatrix}.$$

Cette matrice admet une valeur propre nulle car son rang est inférieur à 3. Afin de trouver le signe des trois valeurs propres restantes, $\Lambda_1, \Lambda_2, \Lambda_3$, on calcule le polynôme caractéristique de $\text{Hess}_{(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)})} f$, noté $\chi(X)$.

On a :

$$\chi(X) = X \left(X^3 + (-a_{11} - a_{22}) X^2 + (-a_{12}^2 - a_{13}^2 - a_{14}^2 + a_{11}a_{22}) X + a_{22} (a_{13}^2 + a_{14}^2) \right).$$

Par les relations coefficients-racines, on a donc :

$$\begin{cases} \Lambda_1 + \Lambda_2 + \Lambda_3 = a_{11} + a_{22} \\ \Lambda_1 \Lambda_2 \Lambda_3 = -a_{22} (a_{13}^2 + a_{14}^2). \end{cases}$$

Or par l'expression de \bar{b}_i donnée en (G.4),

$$a_{11} = - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{\| -z_{t_{k-1}, i} + \bar{\mu}_i \|^2}{\bar{b}_i} < 0 \text{ et } a_{22} = - \frac{1}{\bar{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) < 0.$$

Donc :

$$\begin{cases} \Lambda_1 + \Lambda_2 + \Lambda_3 < 0 \\ \Lambda_1 \Lambda_2 \Lambda_3 > 0. \end{cases}$$

Donc forcément deux valeurs propres sont négatives et une positive.

Ainsi, le point critique $(1, \bar{b}_i, \overline{\mu_i(0)}, \overline{\mu_i(1)})$ est l'argument d'un point "col", ce n'est donc pas l'argument d'un extrémum.

Etude du cas $a_i \neq 1$: On notera $\hat{C} = (\hat{a}_i, \hat{b}_i, \widehat{\mu_i(0)}, \widehat{\mu_i(1)})$ le point critique.

On a :

$$\left\{ \frac{\partial f}{\partial a_i}(\hat{C}) = 0 \right\} \iff \begin{cases} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \langle z_{t_k, i} - z_{t_{k-1}, i} a_i - \mu_i(1 - a_i); -z_{t_{k-1}, i} \rangle \\ + \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) (z_{t_k, i}(0) - z_{t_{k-1}, i}(0) a_i - \mu_i(0)(1 - a_i)) \mu_i(0) \\ + \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) (z_{t_k, i}(1) - z_{t_{k-1}, i}(1) a_i - \mu_i(1)(1 - a_i)) \mu_i(1) = 0 \end{cases}$$

Les deux derniers termes sont nuls par (G.2) donc :

$$\left\{ \frac{\partial f}{\partial a_i}(\widehat{C}) = 0 \iff \begin{cases} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \langle z_{t_k, i} - z_{t_{k-1}, i} a_i - \mu_i(1 - a_i); z_{t_{k-1}, i} \rangle = 0 \\ \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \langle z_{t_k, i} - z_{t_{k-1}, i} a_i; z_{t_{k-1}, i} \rangle \\ - \mu_i(0)(1 - a_i) \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_{k-1}, i}(0) \\ - \mu_i(1)(1 - a_i) \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_{k-1}, i}(1) = 0. \end{cases} \quad (\text{G.5}) \right.$$

Par (G.3), on a :

$$\left\{ \frac{\partial f}{\partial a_i}(\widehat{C}) = 0 \iff \begin{cases} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \langle z_{t_k, i} - z_{t_{k-1}, i} a_i; z_{t_{k-1}, i} \rangle \\ - \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) (z_{t_k, i}(0) - z_{t_{k-1}, i}(0) a_i)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_{k-1}, i}(0) \\ - \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) (z_{t_k, i}(0) - z_{t_{k-1}, i}(0) a_i)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_{k-1}, i}(1) = 0, \end{cases}$$

ce qui nous donne la première coordonnée du point critique $\widehat{a}_i = \frac{n_i}{d_i}$ avec :

$$\begin{aligned} d_i &= \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \right) \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \|z_{t_{k-1}, i}\|^2 \right) \\ &\quad - \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_{k-1}, i}(0) \right)^2 - \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_{k-1}, i}(1) \right)^2 \\ &= \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \right) \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \|z_{t_{k-1}, i}\|^2 - \left\| \sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{Z}, \vartheta) z_{t_{j-1}, i} \right\|^2, \quad (\text{G.6}) \end{aligned}$$

et

$$\begin{aligned}
 n_i &= \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \right) \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \langle z_{t_k, i}; z_{t_{k-1}, i} \rangle \right) \\
 &\quad - \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_k, i}(0) \right) \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_{k-1}, i}(0) \right) \\
 &\quad - \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_k, i}(1) \right) \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_{k-1}, i}(1) \right) \\
 &= \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \right) \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \langle z_{t_{k-1}, i}, z_{t_k, i} \rangle \right) \\
 &\quad - \left\langle \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) z_{t_{k-1}, i}, \sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{Z}, \vartheta) z_{t_j, i} \right\rangle. \quad (\text{G.7})
 \end{aligned}$$

Puis avec (G.3), on trouve $\widehat{\mu}_i(0)$ et $\widehat{\mu}_i(1)$ en fonction de \widehat{a}_i :

$$\left\{ \begin{array}{l}
 \widehat{\mu}_i(0) = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) (z_{t_k, i}(0) - z_{t_{k-1}, i}(0) \widehat{a}_i)}{(1 - \widehat{a}_i) \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)} \\
 \widehat{\mu}_i(1) = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) (z_{t_k, i}(1) - z_{t_{k-1}, i}(1) \widehat{a}_i)}{(1 - \widehat{a}_i) \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)}.
 \end{array} \right.$$

Pour finir, on annule la dernière dérivée et on trouve \widehat{b}_i en fonction de \widehat{a}_i et $\widehat{\mu}_i = \begin{pmatrix} \widehat{\mu}_i(0) \\ \widehat{\mu}_i(1) \end{pmatrix}$:

$$\left\{ \frac{\partial f}{\partial b_i}(\widehat{C}) = 0 \iff \widehat{b}_i = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} \widehat{a}_i - \widehat{\mu}_i(1 - \widehat{a}_i)\|^2}{2 \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)} \geq 0. \right. \quad (\text{G.8})$$

Montrons que le point critique $(\widehat{a}_i, \widehat{b}_i, \widehat{\mu}_i(0), \widehat{\mu}_i(1))$ que nous venons de trouver est un maximum local de la fonction f . Pour cela, on calcule les dérivées secondes

de cette fonction f et on les évalue au point critique \widehat{C} . On a :

$$\frac{\partial^2 f}{\partial a_i^2}(\widehat{C}) = - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{\| -z_{t_{k-1},i} + \widehat{\mu}_i \|^2}{\widehat{b}_i} =: a_{11},$$

et :

$$\frac{\partial^2 f}{\partial b_i^2}(\widehat{C}) = \frac{1}{\widehat{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) - \frac{1}{\widehat{b}_i^3} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \| z_{t_k,i} - z_{t_{k-1},i} \widehat{a}_i - \widehat{\mu}_i (1 - \widehat{a}_i) \|^2.$$

En utilisant l'expression de \widehat{b}_i donnée en (G.8), on a :

$$\begin{aligned} \frac{\partial^2 f}{\partial b_i^2}(\widehat{C}) &= \frac{1}{\widehat{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) - \frac{1}{\widehat{b}_i^3} 2 \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \right) \widehat{b}_i \\ &= - \frac{1}{\widehat{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) =: a_{22}. \end{aligned}$$

On a aussi :

$$\frac{\partial^2 f}{\partial \mu_i(0)^2}(\widehat{C}) = - \frac{(1 - \widehat{a}_i)^2}{\widehat{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) =: a_{33}.$$

De la même manière, on obtient que :

$$\frac{\partial^2 f}{\partial \mu_i(1)^2}(\widehat{C}) = - \frac{(1 - \widehat{a}_i)^2}{\widehat{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) = a_{33}.$$

On a ensuite :

$$\frac{\partial^2 f}{\partial a_i \partial b_i}(\widehat{C}) = \frac{1}{\widehat{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \langle z_{t_k,i} - z_{t_{k-1},i} \widehat{a}_i - \widehat{\mu}_i (1 - \widehat{a}_i); -z_{t_{k-1},i} + \widehat{\mu}_i \rangle.$$

On retrouve alors grâce à l'expression (G.5), et au fait que $\frac{\partial f}{\partial a_i}$ s'annule en le point critique $(\widehat{a}_i, \widehat{b}_i, \widehat{\mu}_i(0), \widehat{\mu}_i(1))$, que :

$$\frac{\partial^2 f}{\partial a_i \partial b_i}(\widehat{C}) = 0.$$

et :

$$\begin{aligned} \frac{\partial^2 f}{\partial a_i \partial \mu_i(0)}(\widehat{C}) &= \frac{\widehat{a}_i - 1}{\widehat{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)(-z_{t_{k-1},i}(0) + \widehat{\mu}_i(0)) \\ &\quad - \frac{1}{\widehat{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)(z_{t_k,i}(0) - z_{t_{k-1},i}(0)\widehat{a}_i - \widehat{\mu}_i(0)(1 - \widehat{a}_i))(\widehat{a}_i - 1). \end{aligned}$$

Le premier terme de cette expression est nul par (G.2), donc on a :

$$\frac{\partial^2 f}{\partial a_i \partial \mu_i(0)}(\widehat{C}) = -\frac{(\widehat{a}_i - 1)}{\widehat{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)(-z_{t_{k-1},i}(0) + \widehat{\mu}_i(0)) =: a_{13}.$$

De la même manière on obtient :

$$\frac{\partial^2 f}{\partial a_i \partial \mu_i(1)}(\widehat{C}) = -\frac{(\widehat{a}_i - 1)}{\widehat{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)(-z_{t_{k-1},i}(0) + \widehat{\mu}_i(1)) =: a_{14}.$$

Aussi, on a par (G.2) :

$$\frac{\partial^2 f}{\partial b_i \partial \mu_i(0)}(\widehat{C}) = \frac{1}{\widehat{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)(z_{t_k,i}(0) - z_{t_{k-1},i}(0)a_i - \mu_i(0)(1 - a_i))(a_i - 1) = 0,$$

et de même :

$$\frac{\partial^2 f}{\partial b_i \partial \mu_i(1)}(\widehat{C}) = 0.$$

Enfin :

$$\frac{\partial^2 f}{\partial \mu_i(0) \partial \mu_i(1)}(\widehat{C}) = 0.$$

La matrice hessienne de f en le point critique $(\widehat{a}_i, \widehat{b}_i, \widehat{\mu}_i(0), \widehat{\mu}_i(1))$ est donc :

$$\text{Hess}_{(\widehat{a}_i, \widehat{b}_i, \widehat{\mu}_i(0), \widehat{\mu}_i(1))} f = \begin{pmatrix} a_{11} & 0 & a_{13} & a_{14} \\ 0 & a_{22} & 0 & 0 \\ a_{13} & 0 & a_{33} & 0 \\ a_{14} & 0 & 0 & a_{33} \end{pmatrix}.$$

Afin de montrer que le point critique $(\widehat{a}_i, \widehat{b}_i, \widehat{\mu}_i(0), \widehat{\mu}_i(1))$ est l'argument d'un maximum local pour f , on va montrer que $\text{Hess}_{(\widehat{a}_i, \widehat{b}_i, \widehat{\mu}_i(0), \widehat{\mu}_i(1))} f$ est définie négative, c'est-à-dire que $-\text{Hess}_{(\widehat{a}_i, \widehat{b}_i, \widehat{\mu}_i(0), \widehat{\mu}_i(1))} f$ est définie positive.

Pour cela, on utilise le critère de Sylvester, en montrant que ses mineurs principaux sont strictement positifs. On doit donc montrer que :

- (a) : $-a_{11} > 0$,
- (b) : $a_{11}a_{22} > 0$,
- (c) : $-a_{22}(a_{11}a_{33} - a_{13}^2) > 0$,
- (d) : $a_{22}a_{33}(a_{11}a_{33} - a_{13}^2 - a_{14}^2) > 0$.

Montrons que ces 4 conditions sont vérifiées. Tout d'abord, par la définition (5.4) de $\gamma_k^i(j, \vartheta)$, ce dernier est toujours positif. Ainsi, on remarque que :

$$\widehat{b}_i = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} \widehat{a}_i - \widehat{\mu}_i(1 - \widehat{a}_i)\|^2}{2 \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)} > 0.$$

Ainsi, on montre que :

$$(a) \text{ est vérifié : } -a_{11} = \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \frac{\| -z_{t_{k-1}, i} + \widehat{\mu}_i \|^2}{\widehat{b}_i} > 0.$$

$$(b) \text{ est vérifié : } -a_{22} = \frac{1}{\widehat{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) > 0 \text{ donc } a_{11}a_{22} > 0.$$

Pour (d) on a :

$$\begin{aligned} a_{11}a_{33} - a_{13}^2 - a_{14}^2 &= (-a_{11})(-a_{33}) - a_{13}^2 - a_{14}^2 \\ &= \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \frac{\| -z_{t_{k-1}, i} + \widehat{\mu}_i \|^2}{\widehat{b}_i} \right) \left(\frac{(1 - \widehat{a}_i)^2}{\widehat{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \right) \\ &\quad - \left(\frac{(\widehat{a}_i - 1)}{\widehat{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) (-z_{t_{k-1}, i}(0) + \widehat{\mu}_i(0)) \right)^2 \\ &\quad - \left(\frac{(\widehat{a}_i - 1)}{\widehat{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) (-z_{t_{k-1}, i}(1) + \widehat{\mu}_i(1)) \right)^2 \end{aligned}$$

On fabrique des coefficients à valeurs dans $[0, 1]$ en factorisant par $\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)$:

$$\begin{aligned}
 a_{11}a_{33} - a_{13}^2 - a_{14}^2 &= \frac{(\widehat{a}_i - 1)^2}{\widehat{b}_i^2} \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \right)^2 \\
 &\times \left[\sum_{k=i_b+1}^{i_d} \frac{\gamma_k^i(\mathfrak{z}, \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)} \left(-z_{t_{k-1}, i}(0) + \widehat{\mu}_i(0) \right)^2 \right. \\
 &+ \sum_{k=i_b+1}^{i_d} \frac{\gamma_k^i(\mathfrak{z}, \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)} \left(-z_{t_{k-1}, i}(1) + \widehat{\mu}_i(1) \right)^2 \\
 &- \left(\sum_{k=i_b+1}^{i_d} \frac{\gamma_k^i(\mathfrak{z}, \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)} \left(-z_{t_{k-1}, i}(0) + \widehat{\mu}_i(0) \right) \right)^2 \\
 &\left. - \left(\sum_{k=i_b+1}^{i_d} \frac{\gamma_k^i(\mathfrak{z}, \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)} \left(-z_{t_{k-1}, i}(1) + \widehat{\mu}_i(1) \right) \right)^2 \right].
 \end{aligned}$$

Par l'inégalité de Jensen appliquée à fonction carrée et aux coefficients

$$\left(\frac{\gamma_k^i(\mathfrak{z}, \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)} \right)_{k \in \llbracket i_b+1, i_d \rrbracket}$$

dont la somme vaut 1, on obtient que :

$$\begin{aligned}
 \sum_{k=i_b+1}^{i_d} \frac{\gamma_k^i(\mathfrak{z}, \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)} \left(-z_{t_{k-1}, i}(0) + \widehat{\mu}_i(0) \right)^2 \\
 \geq \left(\sum_{k=i_b+1}^{i_d} \frac{\gamma_k^i(\mathfrak{z}, \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)} \left(-z_{t_{k-1}, i}(0) + \widehat{\mu}_i(0) \right) \right)^2
 \end{aligned}$$

et

$$\begin{aligned}
 \sum_{k=i_b+1}^{i_d} \frac{\gamma_k^i(\mathfrak{z}, \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)} \left(-z_{t_{k-1}, i}(1) + \widehat{\mu}_i(1) \right)^2 \\
 \geq \left(\sum_{k=i_b+1}^{i_d} \frac{\gamma_k^i(\mathfrak{z}, \vartheta)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta)} \left(-z_{t_{k-1}, i}(1) + \widehat{\mu}_i(1) \right) \right)^2.
 \end{aligned}$$

Donc $a_{11}a_{33} - a_{13}^2 - a_{14}^2 \geq 0$.

Cela implique que $a_{11}a_{33} - a_{13}^2 \geq 0$, et comme $-a_{22} > 0$, on a aussi montré (c).

Ainsi, $-\text{Hess}_{(\widehat{a}_i, \widehat{b}_i, \widehat{\mu}_i(0), \widehat{\mu}_i(1))} f$ est définie positive, donc $\text{Hess}_{(\widehat{a}_i, \widehat{b}_i, \widehat{\mu}_i(0), \widehat{\mu}_i(1))} f$ est définie négative donc le point critique $(\widehat{a}_i, \widehat{b}_i, \widehat{\mu}_i(0), \widehat{\mu}_i(1))$ est un maximum local de f . L'argument du maximum que l'on recherche doit vérifier $\widehat{b}_i > 0$ et $\widehat{a}_i > 0$ afin de pouvoir appliquer le changement de variable inverse pour retrouver les inconnues initiales :

$$\begin{aligned}\widehat{\lambda}_i &= -\frac{1}{\Delta} \ln(\widehat{a}_i), \\ \widehat{\sigma}_{i,3}^2 &= \frac{b_i}{v(\widehat{\lambda}_i)}.\end{aligned}$$

Si $\widehat{a}_i \leq 0$, regardons ce qu'il se passe sur "le bord $a_i = 0$ ".

Si $a_i = 0$ (ou plutôt $a_i \rightarrow 0^+$ car 0 n'est pas atteint), l'expression de f devient :

$$f(0, b_i, \mu_i(0), \mu_i(1)) = - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln(2\pi b_i) - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{\|z_{t_k, i} - \mu_i\|^2}{2b_i}.$$

Étudions les extrema de cette fonction des variables b_i , $\mu_i(0)$, et $\mu_i(1)$, que l'on note φ :

$$\varphi : (b_i, \mu_i(0), \mu_i(1)) \longmapsto - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln(2\pi b_i) - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{\|z_{t_k, i} - \mu_i\|^2}{2b_i}.$$

Calculons ses dérivées :

$$\left\{ \begin{aligned}\frac{\partial \varphi}{\partial \mu_i(0)}(b_i, \mu_i(0), \mu_i(1)) &= \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{z_{t_k, i}(0) - \mu_i(0)}{b_i} \\ \frac{\partial \varphi}{\partial \mu_i(1)}(b_i, \mu_i(0), \mu_i(1)) &= \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \frac{z_{t_k, i}(1) - \mu_i(1)}{b_i} \\ \frac{\partial \varphi}{\partial b_i}(b_i, \mu_i(0), \mu_i(1)) &= -\frac{1}{b_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) + \frac{1}{2b_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - \mu_i\|^2\end{aligned}\right.$$

et on résout le système suivant d'inconnus \widetilde{b}_i , $\widetilde{\mu}_i(0)$, et $\widetilde{\mu}_i(1)$ pour chercher le point critique de φ .

$$\left\{ \begin{array}{l} \frac{\partial \varphi}{\partial \mu_i(0)}(\tilde{b}_i, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)}) = 0 \\ \frac{\partial \varphi}{\partial \mu_i(1)}(\tilde{b}_i, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)}) = 0 \\ \frac{\partial \varphi}{\partial \tilde{b}_i}(\tilde{b}_i, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)}) = 0 \end{array} \right\} \iff \left\{ \begin{array}{l} \widetilde{\mu_i(0)} = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) z_{t_k, i}(0)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta)} \\ \widetilde{\mu_i(1)} = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) z_{t_k, i}(1)}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta)} \\ \tilde{b}_i = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - \widetilde{\mu_i}\|^2}{2 \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta)} \end{array} \right. \quad (\text{G.9})$$

Montrons que ce point critique $(\tilde{b}_i, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)})$ est l'argument d'un maximum local de la fonction φ . Pour cela, on calcule les dérivées secondes de cette fonction φ et on les évalue au point $(\tilde{b}_i, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)})$:

$$\frac{\partial^2 \varphi}{\partial \tilde{b}_i^2}(\tilde{b}_i, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)}) = \frac{1}{\tilde{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) - \frac{1}{\tilde{b}_i^3} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - \widetilde{\mu_i}\|^2.$$

Ainsi en utilisant l'expression de \tilde{b}_i donnée en (G.9), on a :

$$\begin{aligned} \frac{\partial^2 \varphi}{\partial \tilde{b}_i^2}(\tilde{b}_i, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)}) &= \frac{1}{\tilde{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) - \frac{1}{\tilde{b}_i^3} 2 \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \right) \tilde{b}_i \\ &= -\frac{1}{\tilde{b}_i^2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) =: \widetilde{a_{22}}. \end{aligned}$$

On a aussi :

$$\frac{\partial^2 \varphi}{\partial \mu_i(0)^2}(\tilde{b}_i, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)}) = -\frac{1}{\tilde{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) =: \widetilde{a_{33}}.$$

De la même manière, on obtient que :

$$\frac{\partial^2 \varphi}{\partial \mu_i(1)^2}(\tilde{b}_i, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)}) = -\frac{1}{\tilde{b}_i} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) = \widetilde{a_{33}}.$$

Enfin, on a par l'expression de $\widetilde{\mu_i(0)}$

La matrice hessienne de φ en le point critique $(\widetilde{b_i}, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)})$ est donc :

$$\text{Hess}_{(\widetilde{b_i}, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)})} \varphi = \begin{pmatrix} a_{22} & 0 & 0 \\ 0 & a_{33} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}.$$

Or comme $\widetilde{a_{22}}$ et $\widetilde{a_{33}}$ sont strictement négatifs, les trois valeurs propres de $\text{Hess}_{(\widetilde{b_i}, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)})} \varphi$ sont négatives donc φ admet un maximum local en $(\widetilde{b_i}, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)})$. Ainsi, $(b_i, \mu_i(0), \mu_i(1)) \mapsto f(0, b_i, \mu_i(0), \mu_i(1))$ admet une borne supérieure locale en $(\widetilde{b_i}, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)})$ car ce point n'est pas atteint puisque $a_i = e^{-\lambda_i \Delta} \neq 0$.

Ainsi, pour conclure notre étude précédente, si $\widehat{a_i} > 0$ alors f a un maximum local, atteint en $(\widehat{a_i}, \widehat{b_i}, \widehat{\mu_i(0)}, \widehat{\mu_i(1)})$ et si $\widehat{a_i} \leq 0$ alors f admet une borne supérieure locale sur le bord $a_i = 0$. Ce sup vaut $f(0, \widetilde{b_i}, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)})$.

On remarque que l'expression des variables $\widehat{a_i}, \widehat{b_i}, \widehat{\mu_i(0)}, \widehat{\mu_i(1)}$ s'applique aussi avec $a_i = 0$: en effet, si on remplace $\widehat{a_i}$ par 0 dans les variables $(\widehat{a_i}, \widehat{b_i}, \widehat{\mu_i(0)}, \widehat{\mu_i(1)})$, on trouve $(0, \widetilde{b_i}, \widetilde{\mu_i(0)}, \widetilde{\mu_i(1)})$.

Montrons maintenant que le point $(a_i^{\max}, b_i^{\max}, \mu_i^{\max}(0), \mu_i^{\max}(1))$ est l'argument du maximum global de la fonction f , où :

$$a_i^{\max} = \begin{cases} \frac{n_i}{d_i} & \text{si } n_i > 0 \\ 0 & \text{si } n_i \leq 0 \end{cases}$$

où n_i et d_i sont définis en (G.7) et (G.6), et :

$$\mu_i^{\max} = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)(z_{t_k, i} - z_{t_{k-1}, i} a_i^{\max})}{(1 - a_i^{\max}) \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)}$$

puis :

$$b_i^{\max} = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} a_i^{\max} - \mu_i^{\max}(1 - a_i^{\max})\|^2}{2 \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)}.$$

$$\text{On note } K : \begin{cases} \mathbb{R}_+^* \times \mathbb{R}^2 & \rightarrow \mathbb{R} \\ (a, \mu) & \mapsto \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} a - \mu(1 - a)\|^2. \end{cases}$$

On doit donc prouver que :

$$\forall a > 0, \forall b \geq 0, \forall \mu \in \mathbb{R}^2, f(a, b, \mu(0), \mu(1)) \leq f(a_i^{\max}, b_i^{\max}, \mu_i^{\max}(0), \mu_i^{\max}(1)).$$

Simplifions $f(a_i^{\max}, b_i^{\max}, \mu_i^{\max}(0), \mu_i^{\max}(1))$:

On remarque que :

$$b_i^{\max} = \frac{K(a_i^{\max}, \mu_i^{\max})}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)}, \quad (\text{G.10})$$

d'où :

$$\begin{aligned} f(a_i^{\max}, b_i^{\max}, \mu_i^{\max}(0), \mu_i^{\max}(1)) &= - \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \ln(2\pi b_i^{\max}) \\ &\quad - \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \frac{\|z_{t_k, i} - z_{t_{k-1}, i} a_i^{\max} - \mu_i(1 - a_i^{\max})\|^2}{2b_i^{\max}} \\ &= - \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \ln(2\pi b_i^{\max}) - \frac{2 \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) b_i^{\max}}{2b_i^{\max}} \\ &= - \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \ln(2\pi b_i^{\max}) - \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta). \end{aligned}$$

Il faut donc prouver que $\forall a > 0, \forall b \geq 0, \forall \mu \in \mathbb{R}^2$:

$$\begin{aligned}
 f(a, b, \mu(0), \mu(1)) &\leq f(a_i^{\max}, b_i^{\max}, \mu_i^{\max}(0), \mu_i^{\max}(1)) \\
 \iff - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln(2\pi b) - \frac{K(a, \mu)}{b} &\leq - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln(2\pi b_i^{\max}) - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \\
 \iff K(a, \mu) &\geq b \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln(2\pi b_i^{\max}) - \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \ln(2\pi b) + \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \right) \\
 \iff K(a, \mu) &\geq b (\ln(b_i^{\max}) - \ln(b) + 1) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta).
 \end{aligned}$$

On étudie la fonction $G : b \mapsto b (\ln(b_i^{\max}) - \ln(b) + 1) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta)$. Sa dérivée est :

$$G'(b) = (\ln(b_i^{\max}) - \ln(b)) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta).$$

Ainsi, on a la tableau de variation suivant :

b	0	b_i^{\max}	$+\infty$
$G'(b)$	+	0	-
G			

donc $\forall b \geq 0, G(b) \leq G(b_i^{\max}) = b_i^{\max} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta)$.

Il reste donc à prouver que :

$$\forall a > 0, \forall \mu \in \mathbb{R}^2, K(a, \mu) \geq G(b_i^{\max}) = b_i^{\max} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta).$$

Étudions la fonction $k : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R} \\ \mu & \mapsto K(a, \mu) \end{cases}$:

$$\begin{aligned}
 k(\mu) &= \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} a - \mu(1-a)\|^2 \\
 &= \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \left(\|z_{t_k, i} - z_{t_{k-1}, i} a\|^2 + (1-a)^2 \|\mu\|^2 - 2 \langle z_{t_k, i} - z_{t_{k-1}, i} a, (1-a)\mu \rangle \right) \\
 &= \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} a\|^2 + (1-a)^2 \|\mu\|^2 \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \\
 &\quad - (1-a) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \langle z_{t_k, i} - z_{t_{k-1}, i} a, \mu \rangle,
 \end{aligned}$$

donc :

$$\begin{aligned}
 k(\mu) &= \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} a\|^2 \\
 &\quad + (1-a)^2 \mu(0)^2 \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) - (1-a)\mu(0) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) (z_{t_k, i}(0) - z_{t_{k-1}, i}(0)a) \\
 &\quad + (1-a)^2 \mu(1)^2 \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) - (1-a)\mu(1) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) (z_{t_k, i}(1) - z_{t_{k-1}, i}(1)a).
 \end{aligned}$$

• **Traisons le cas $a \neq 1$:**

La fonction k est alors la somme d'un polynôme de degré 2 en $\mu(0)$ (avec un coefficient strictement positif devant $\mu(0)^2$) et d'un polynôme de degré 2 en $\mu(1)$ (avec un coefficient strictement positif devant $\mu(1)^2$), donc k atteint son minimum en :

$$\mu^m = \left(\frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) (z_{t_k, i}(0) - z_{t_{k-1}, i}(0)a)}{(1-a) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta)}, \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) (z_{t_k, i}(1) - z_{t_{k-1}, i}(1)a)}{(1-a) \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta)} \right)$$

et :

$$k(\mu^m) = \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \left\| z_{t_k, i} - z_{t_{k-1}, i} a - \frac{\sum_{j=i_b+1}^{i_d} \gamma_j^i(3, \vartheta) (z_{t_j, i} - z_{t_{j-1}, i} a)}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(3, \vartheta)} \right\|^2.$$

Comme $\forall \mu \in \mathbb{R}^2$, $K(a, \mu) \geq K(a, \mu^m)$, il suffit de montrer que

$$\forall a > 0, K(a, \mu^m) \geq G(b_i^{\max}) = b_i^{\max} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta),$$

avec $b_i^{\max} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) = K(a_i^{\max}, \mu_i^{\max})$ par (G.10).

Ainsi, montrons que :

$$\forall a > 0, a \neq 1, K(a, \mu^m) \geq K(a_i^{\max}, \mu_i^{\max}),$$

avec :

$$\begin{aligned} K(a_i^{\max}, \mu_i^{\max}) &= \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} a_i^{\max} - \mu_i^{\max} (1 - a_i^{\max})\|^2 \\ &= \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \left\| z_{t_k, i} - z_{t_{k-1}, i} a_i^{\max} - \frac{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{Z}, \vartheta) (z_{t_j, i} - z_{t_{j-1}, i} a_i^{\max})}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{Z}, \vartheta)} \right\|^2. \end{aligned}$$

Si on note :

$$H : \begin{cases} \mathbb{R}_+^* & \rightarrow \mathbb{R} \\ a & \mapsto \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \left\| z_{t_k, i} - z_{t_{k-1}, i} a - \frac{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{Z}, \vartheta) (z_{t_j, i} - z_{t_{j-1}, i} a)}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{Z}, \vartheta)} \right\|^2, \end{cases}$$

cela revient à montrer que :

$$\forall a > 0, a \neq 1, H(a) \geq H(a_i^{\max}). \quad (\text{G.11})$$

On a :

$$\begin{aligned}
 H(a) = & \frac{a^2}{2} \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_{k-1}, i}\|^2 - \frac{\left\| \sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta) z_{t_{j-1}, i} \right\|^2}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta)} \right) \\
 & + a \left(- \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \langle z_{t_{k-1}, i}, z_{t_k, i} \rangle + \left\langle \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) z_{t_{k-1}, i}, \frac{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta) z_{t_j, i}}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta)} \right\rangle \right) \\
 & + \frac{1}{2} \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_k, i}\|^2 - \frac{\left\| \sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta) z_{t_j, i} \right\|^2}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta)} \right).
 \end{aligned}$$

H est donc un polynôme de degré 2 avec un coefficient strictement positif (grâce à l'inégalité de Jensen) devant a^2 . Ainsi, H atteint son minimum en :

$$a^m = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \langle z_{t_{k-1}, i}, z_{t_k, i} \rangle - \left\langle \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) z_{t_{k-1}, i}, \frac{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta) z_{t_j, i}}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta)} \right\rangle}{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_{k-1}, i}\|^2 - \frac{\left\| \sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta) z_{t_{j-1}, i} \right\|^2}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta)}}.$$

On remarque que $a^m = a_i^{\max}$ dans le cas où $n_i > 0 \Leftrightarrow a_i^{\max} > 0$. Ainsi :

— Si $a^m > 0$, alors $a_i^{\max} = a^m > 0$ donc

$$H(a) \geq H(a^m) = H(a_i^{\max}), \forall a > 0, a \neq 1.$$

— Si $a^m \leq 0$ alors $H(a) \geq H(0), \forall a > 0, a \neq 1$, car H est croissante sur

$[a^m, +\infty[$ avec $a^m \leq 0$, et dans ce cas $a_i^{\max} = 0$ donc

$$H(a) \geq H(0) = H(a_i^{\max}), \forall a > 0, a \neq 1.$$

On a donc bien montré que $H(a) \geq H(a_i^{\max})$ lorsque $a_i^{\max} > 0$ et lorsque $a_i^{\max} = 0$, on a donc prouvé (G.11), donc :

$$K(a, \mu) \geq K(a_i^{\max}, \mu_i^{\max}), \forall a > 0, a \neq 1, \forall \mu \in \mathbb{R}^2.$$

• **Traitons maintenant le cas $a = 1$:**

Il nous reste à prouver que :

$$\forall \mu \in \mathbb{R}^2, K(1, \mu) \geq G(b_i^{\max}) = b_i^{\max} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) = K(a_i^{\max}, \mu_i^{\max}) = H(a_i^{\max}).$$

On a :

$$K(1, \mu) = \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2.$$

Montrons que :

$$\frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2 \geq H(a_i^{\max}).$$

On calcule $H(1)$:

$$\begin{aligned} H(1) = & \frac{1}{2} \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_{k-1}, i}\|^2 - \frac{\left\| \sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta) z_{t_{j-1}, i} \right\|^2}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta)} \right) \\ & + \left(- \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \langle z_{t_{k-1}, i}, z_{t_k, i} \rangle + \left\langle \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) z_{t_{k-1}, i}, \frac{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta) z_{t_j, i}}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta)} \right\rangle \right) \\ & + \frac{1}{2} \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_k, i}\|^2 - \frac{\left\| \sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta) z_{t_j, i} \right\|^2}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta)} \right), \end{aligned}$$

donc :

$$H(1) = \frac{1}{2} \left(\sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2 - \frac{\left\| \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) z_{t_{k-1}, i} - \sum_{j=i_b+1}^{i_d} \gamma_j^i(3, \vartheta) z_{t_j, i} \right\|^2}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(3, \vartheta)} \right),$$

donc :

$$H(1) \leq \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2.$$

De plus, $H(a_i^{\max}) \leq H(1)$ car si $a_i^{\max} > 0$, alors $a_i^{\max} = a^m$ est l'argument du minimum de H et si $a_i^{\max} = 0$ alors $a^m \leq 0$ et donc H est croissante sur \mathbb{R}_+ .

Donc $H(a_i^{\max}) \leq H(1) \leq \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2$, ie :

$$H(a_i^{\max}) \leq \frac{1}{2} \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2 = K(1, \mu), \forall \mu \in \mathbb{R}^2.$$

Ainsi, on a montré que :

$$\forall a > 0, \forall \mu \in \mathbb{R}^2, K(a, \mu) \geq K(a_i^{\max}, \mu_i^{\max}).$$

Donc :

$$\forall a > 0, \forall b \geq 0, \forall \mu \in \mathbb{R}^2, f(a, b, \mu(0), \mu(1)) \leq f(a_i^{\max}, b_i^{\max}, \mu_i^{\max}(0), \mu_i^{\max}(1)).$$

Donc $(a_i^{\max}, b_i^{\max}, \mu_i^{\max}(0), \mu_i^{\max}(1))$ est l'argument du maximum global de f (atteint ou non).

Pour retrouver l'argument du maximum avec les inconnues initiales, on applique le changement de variable inverse.

La dérivée de cette fonction s'annule en :

$$\sigma_b^2 = \frac{1}{2\Delta} \frac{\sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i}\|^2}{\sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta)}.$$

Ensuite, pour trouver l'argument du maximum de Q selon les variables $(\rho, \eta_i, \sigma_r^2) \in \mathbb{R}_+^* \times [0, 2\pi[\times \mathbb{R}_+^*$, on cherche pour tout $i \in \llbracket 1, s \rrbracket$ le maximum de la fonction :

$$(\rho, \eta_i, \sigma_r^2) \mapsto \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(2, \vartheta) \ln \left(b^j \left(z_{t_k, j}, z_{t_{k-1}, j}, 2, (\rho, \eta_j, \sigma_r^2) \right) \right).$$

Comme précédemment, par la forme de $b^i(\cdot, \cdot, 2, \cdot)$ donnée dans la proposition 10 du chapitre 5, on a :

$$\begin{aligned} & \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(2, \vartheta) \ln \left(b^j \left(z_{t_k, j}, z_{t_{k-1}, j}, 2, (\rho, \eta_j, \sigma_r^2) \right) \right) \\ &= - \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(2, \vartheta) \ln \left(2\pi\sigma_r^2\Delta \right) - \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(2, \vartheta) \frac{\left\| z_{t_k, j} - z_{t_{k-1}, j} - \rho \begin{pmatrix} \cos(\eta_j) \\ \sin(\eta_j) \end{pmatrix} \Delta \right\|^2}{2\sigma_r^2\Delta}. \end{aligned}$$

On calcule puis on annule les dérivées par rapport aux variables ρ, η_i et σ_r^2 et on obtient :

$$\begin{aligned} \eta_i &= \arctan \left(\frac{\sum_{k=i_b+1}^{i_d} (z_{t_k, i}(1) - z_{t_{k-1}, i}(1)) \gamma_k^i(2, \vartheta)}{\sum_{k=i_b+1}^{i_d} (z_{t_k, i}(0) - z_{t_{k-1}, i}(0)) \gamma_k^i(2, \vartheta)} \right), \\ \rho &= \frac{1}{\Delta} \frac{\sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \left\langle \begin{pmatrix} \cos(\eta_j) \\ \sin(\eta_j) \end{pmatrix}, z_{t_k, j} - z_{t_{k-1}, j} \right\rangle \gamma_k^j(2, \vartheta)}{\sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(2, \vartheta)}, \\ \sigma_r^2 &= \frac{1}{2\Delta} \frac{\sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(2, \vartheta) \left\| z_{t_k, j} - z_{t_{k-1}, j} - \rho \begin{pmatrix} \cos(\eta_j) \\ \sin(\eta_j) \end{pmatrix} \Delta \right\|^2}{\sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(2, \vartheta)}. \end{aligned}$$

Pour trouver, pour tout $i \in \llbracket 1, s \rrbracket$, l'argument du maximum de Q selon les variables $(\lambda, (\mu_i(0), \mu_i(1)), \sigma_v^2) \in \mathbb{R}_+ \times \Lambda \times \mathbb{R}_+^*$, on cherche le maximum de la fonction :

$$f : (\lambda, \mu_i(0), \mu_i(1), \sigma_v^2) \mapsto \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(\mathfrak{z}, \vartheta) \ln \left(b^j \left(z_{t_k,j}, z_{t_{k-1},j}, \mathfrak{z}, (\lambda, \mu_j(0), \mu_j(1), \sigma_v^2) \right) \right).$$

On a :

$$\begin{aligned} & \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(\mathfrak{z}, \vartheta) \ln \left(b^j \left(z_{t_k,j}, z_{t_{k-1},j}, \mathfrak{z}, (\lambda, \mu_j(0), \mu_j(1), \sigma_v^2) \right) \right) \\ &= - \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(\mathfrak{z}, \vartheta) \ln \left(2\pi\sigma_v^2 v(\lambda) \right) - \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(\mathfrak{z}, \vartheta) \frac{\|z_{t_k,j} - z_{t_{k-1},j} e^{-\lambda\Delta} - \mu_j(1 - e^{-\lambda\Delta})\|^2}{2\sigma_v^2 v(\lambda)}. \end{aligned}$$

On fait le même changement de variable :

$$\begin{cases} a = e^{-\lambda\Delta} \\ b = \sigma_v^2 v(\lambda) \\ \mu_i(0) = \mu_i(0) \\ \mu_i(1) = \mu_i(1) \end{cases}$$

et on trouve le point critique :

$$a = \frac{n}{d}$$

avec :

$$n = \sum_{i=1}^s \left[\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \langle z_{t_{k-1},i}, z_{t_k,i} \rangle - \left\langle \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) z_{t_{k-1},i}, \frac{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta) z_{t_j,i}}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta)} \right\rangle \right],$$

et

$$d = \sum_{i=1}^s \left[\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{z}, \vartheta) \|z_{t_{k-1},i}\|^2 - \frac{\left\| \sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta) z_{t_{j-1},i} \right\|^2}{\sum_{j=i_b+1}^{i_d} \gamma_j^i(\mathfrak{z}, \vartheta)} \right].$$

Puis, comme précédemment, avec (G.3), on trouve $\mu_i(0)$ et $\mu_i(1)$ en fonction de a .

On a :

$$\mu_i(0) = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)(z_{t_k, i}(0) - z_{t_{k-1}, i}(0)a)}{(1-a) \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)},$$

et :

$$\mu_i(1) = \frac{\sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)(z_{t_k, i}(1) - z_{t_{k-1}, i}(1)a)}{(1-a) \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)}.$$

Pour finir, on résout la dernière équation du système en fonction de a et de $\mu_i = \begin{pmatrix} \mu_i(0) \\ \mu_i(1) \end{pmatrix}$ pour tout $i \in \llbracket 1, s \rrbracket$:

$$\begin{cases} \frac{\partial f}{\partial b}(a, b, \mu_i(0), \mu_i(1)) = 0 \\ \Leftrightarrow \begin{cases} b = \frac{\sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta) \|z_{t_k, i} - z_{t_{k-1}, i} a - \mu_i(1-a)\|^2}{2 \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(\mathfrak{Z}, \vartheta)}. \end{cases} \end{cases}$$

Remarque : On procède comme dans la preuve précédente pour montrer que ce maximum est atteint si $a > 0$ et que sinon, on obtient une borne supérieure, en le même point dans lequel on remplace a par 0. Puis, on utilise les mêmes techniques de calcul que précédemment pour montrer que ce maximum local est en fait un maximum global.

G.4 Preuve de la proposition 14 du chapitre 5

Dans le cas où $\sigma_b = \sigma_r = \sigma_v = \sigma$, on trouve de la même manière que dans la preuve précédente les arguments du maximum pour les variables $\tau, p_b, p_r, p_v, \lambda, \rho, \eta_i, \mu_i$, puis on cherche ensuite à exprimer le maximum de la fonction :

$$f : \sigma^2 \mapsto \sum_{i=1}^3 \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(i, \vartheta) \ln \left(b^j \left(z_{t_k,j}, z_{t_{k-1},j}, i, (\sigma^2) \right) \right).$$

On a :

$$\begin{aligned} f(\sigma^2) = & - \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(1, \vartheta) \ln(2\pi\sigma^2\Delta) - \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(1, \vartheta) \frac{\|z_{t_k,j} - z_{t_{k-1},j}\|^2}{2\sigma^2\Delta} \\ & - \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(2, \vartheta) \ln(2\pi\sigma^2\Delta) - \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(2, \vartheta) \frac{\left\| z_{t_k,j} - z_{t_{k-1},j} - \rho \begin{pmatrix} \cos(\eta_j) \\ \sin(\eta_j) \end{pmatrix} \Delta \right\|^2}{2\sigma^2\Delta} \\ & - \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(3, \vartheta) \ln(2\pi\sigma^2v(\lambda)) - \sum_{k=j_b+1}^{j_d} \gamma_k^j(3, \vartheta) \frac{\|z_{t_k,j} - z_{t_{k-1},j}e^{-\lambda\Delta} - \mu_j(1 - e^{-\lambda\Delta})\|^2}{2\sigma^2v(\lambda)}. \end{aligned}$$

La dérivée de cette fonction s'annule en :

$$\sigma^2 = \frac{n_\sigma}{\sum_{j=1}^s \sum_{k=j_b+1}^{j_d} (\gamma_k^j(1, \vartheta) + \gamma_k^j(2, \vartheta) + \gamma_k^j(3, \vartheta))},$$

où :

$$\begin{aligned} n_\sigma = & \frac{1}{2\Delta} \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(1, \vartheta) \|z_{t_k,i} - z_{t_{k-1},i}\|^2 \\ & + \frac{1}{2\Delta} \sum_{j=1}^s \sum_{k=j_b+1}^{j_d} \gamma_k^j(2, \vartheta) \left\| z_{t_k,j} - z_{t_{k-1},j} - \rho \begin{pmatrix} \cos(\eta_j) \\ \sin(\eta_j) \end{pmatrix} \Delta \right\|^2 \\ & + \frac{1}{2v(\lambda)} \sum_{i=1}^s \sum_{k=i_b+1}^{i_d} \gamma_k^i(3, \vartheta) \|z_{t_k,i} - z_{t_{k-1},i}a - \mu_i(1 - a)\|^2. \end{aligned}$$

Comme précédemment, on montre qu'il s'agit d'un maximum local puis global.

COMPARAISON AVEC LA MÉTHODE DE [MBP⁺15]

Dans son article [MBP⁺15], Monnier présente une méthode aussi basée sur les chaînes de Markov à états cachés, qui a pour but, pour une trajectoire donnée de déterminer quel morceau suit un mouvement brownien et quel morceau suit un mouvement brownien dirigé et de déterminer, pour chacun des mouvements présents le long de la trajectoire ses différents paramètres (le coefficient de diffusion pour un mouvement brownien, le coefficient de diffusion et le drift pour un mouvement brownien dirigé) grâce à une méthode bayésienne basée sur l'algorithme Metropolis MCMC.

Le force de cette méthode est qu'elle prend en argument un nombre entier K_{\max} , qui sera le nombre maximum de mouvements différents qu'on pourra retrouver sur la trajectoire en question et étudie parmi toutes les combinaisons de $K \leq K_{\max}$ mouvements possibles, laquelle est la plus probable pour décrire la trajectoire étudiée. Les K mouvements étudiés sont choisis parmi plusieurs mouvements browniens (notés D) et/ou plusieurs mouvements browniens dirigés (notés DV). Ainsi, par exemple, pour $K_{\max} = 2$, la méthode teste 5 combinaisons possibles : {D}, {DV}, {D,D}, {D, DV}, {DV, DV}.

En revanche, cette méthode ne convient pas pour étudier un ensemble de trajectoires dans lequel toutes les trajectoires suivent toutes les mêmes types de mouvement avec les mêmes paramètres. En effet, lorsque l'on suppose que plusieurs trajectoires suivent des mouvements similaires (browniens avec même coefficient de diffusion, brownien dirigé avec le même coefficient de diffusion et la même norme de drift), alors la méthode cherche à trouver un vecteur de drift commun. Or si le drift des trajectoires, bien qu'ils aient la même norme, ne vont pas dans la même

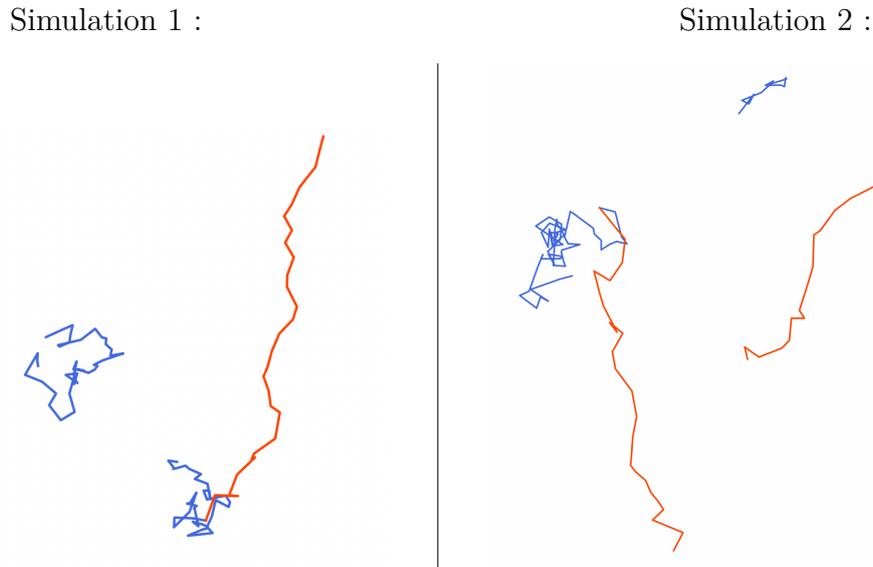


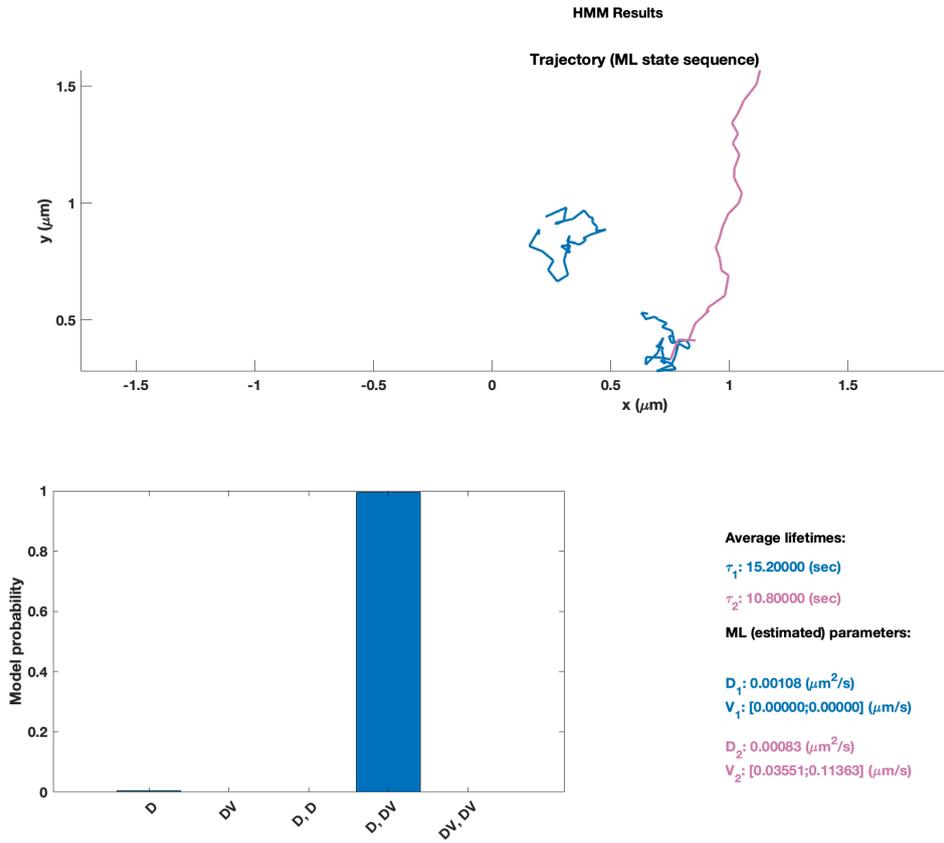
FIGURE H.1 – Deux simulations de trajectoires suivant un mouvement brownien de coefficient de diffusion $\sigma = 0.8$ sur les parties **bleues** et un mouvement brownien dirigé de même coefficient de diffusion et de drift de norme $\|v\| = 3.5$ sur les parties **rouge**.

direction (cas que l'on ne peut pas exclure pour l'étude de données réelles), alors la méthode ne pourra pas trouver de vecteur de drift commun et considérera donc qu'il s'agit d'un unique brownien non dirigé.

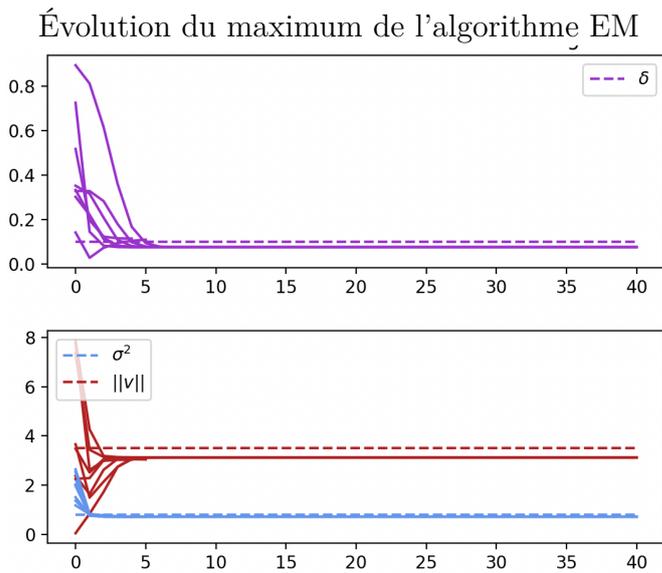
La figure H.1 présente deux simulations : de deux trajectoires pour la Simulation 1 (à gauche) et de trois trajectoires pour la Simulation 2 (à droite). Ces trajectoires sont simulées avec le même coefficient de diffusion pour les parties browniennes et pour les parties browniennes dirigées et la même norme de drift (mais pas le même angle) pour les parties browniennes dirigées.

La figure H.2 regroupe les résultats de l'analyse des deux trajectoires de la Simulation 1 de la figure H.1. En haut, on retrouve les résultats de la méthode Monnier pour $K_{\max} = 2$. L'historgramme en bas à gauche de cette figure indique que parmi toutes les combinaisons possibles d'au plus 2 mouvements, celle qui est la plus probable pour ces deux trajectoires est $\{D, DV\}$ et les deux trajectoires sont ensuite coloriées sachant que c'est cette combinaison qui est choisie (graphique du haut avec en bleu les parties de trajectoires classées D et en rose les parties de trajectoire classées DV). On voit que dans cet exemple la méthode Monnier fonctionne bien.

Méthode Monnier :



Ma méthode :



Trajectoires coloriées avec les résultats finaux de l'algorithme EM

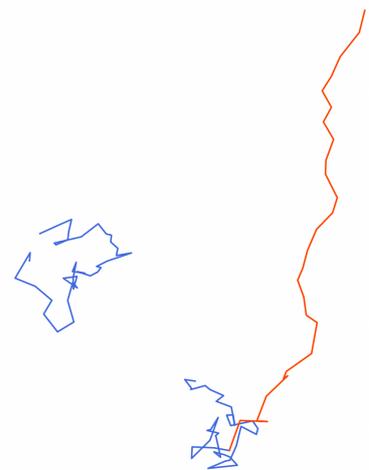


FIGURE H.2 – Résultats de la méthode Monnier (en haut) et de ma méthode (en bas) sur les trajectoires de la Simulation 1.

En bas de la figure H.2, on trouve les résultats de ma méthode pour ces deux mêmes trajectoires, adaptée pour le cas où seuls deux types de mouvements sont possibles, un mouvement brownien de coefficient de diffusion σ et un mouvement dirigé de coefficient de diffusion σ et de drift de norme $\|v\|$. Dans ce cas la matrice de transition a est de taille 2×2 et est donc de la forme

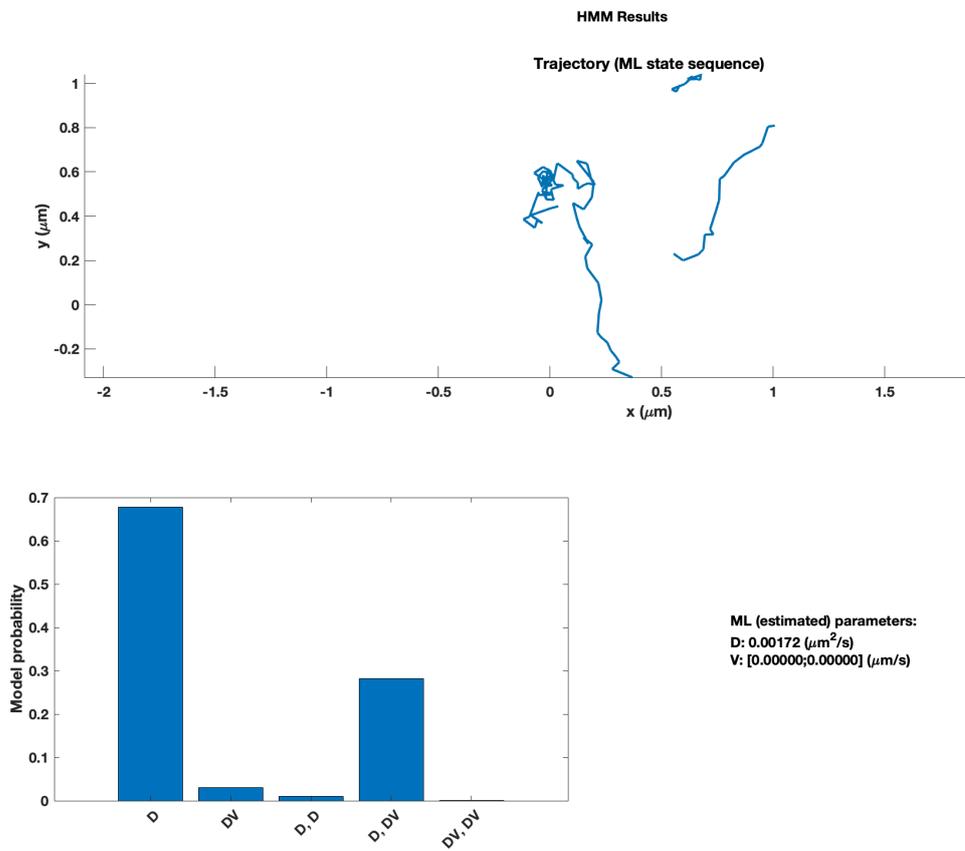
$$\begin{matrix} & \begin{matrix} 1 & 2 \end{matrix} \\ \begin{matrix} 1 \\ 2 \end{matrix} & \begin{pmatrix} e^{-\delta\Delta} & 1 - e^{-\delta\Delta} \\ 1 - e^{-\delta\Delta} & e^{-\delta\Delta} \end{pmatrix} \end{matrix}$$

avec le taux δ comme inconnue. Le graphique à gauche montre la convergence des paramètres δ, σ et $\|v\|$ renvoyés par l’algorithme EM en fonction du nombre d’itérations. Les pointillés désignent les valeurs théoriques de ces différents paramètres, utilisés pour simuler les trajectoires. On voit donc quelque soit les points de départ choisis au hasard, l’algorithme EM converge bien. Ensuite, la figure de droite affiche le résultat de l’algorithme de Viterbi, utilisé avec les paramètres finaux renvoyés par l’algorithme EM. On voit que les deux trajectoires sont correctement classifiées.

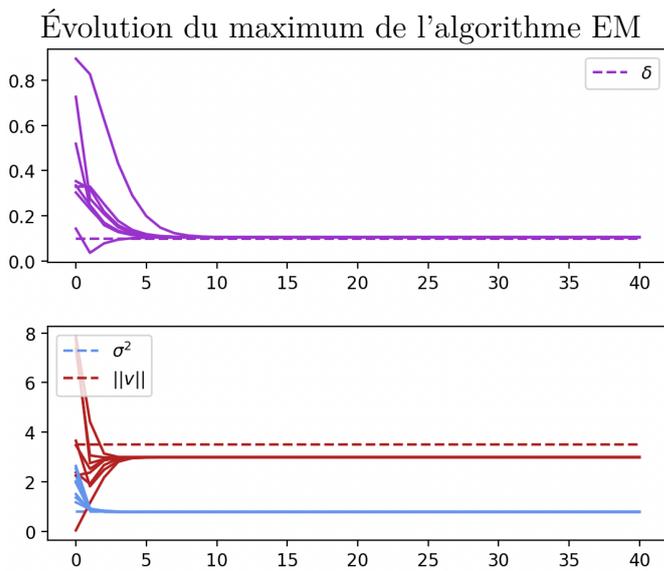
La figure H.3 regroupe les résultats de l’analyse des trois trajectoires de la Simulation 2 de la figure H.1. En haut, on retrouve les résultats de la méthode Monnier pour $K_{\max} = 2$. On voit grâce à l’histogramme en bas à gauche que la méthode Monnier a choisi que la combinaison de mouvements la plus probable pour ces trois trajectoires était $\{D\}$, donc un seul brownien. Or, comme on le voit sur la figure H.1, deux parties de ces trois trajectoires sont issues d’un mouvement brownien dirigé. Ceci s’explique car la méthode de Monnier est faite pour trouver les coordonnées x et y du vecteur drift du mouvement brownien dirigé (s’il considère qu’il y en a un). Or ici, bien que les parties de trajectoires rouges aient la même norme de drift, leur vecteur drift ne va pas dans la même direction. Donc l’algorithme ne peut pas leur trouver de vecteur drift commun. C’est ce qui explique aussi pourquoi dans le jeu de trajectoires de la Simulation 1, cela fonctionnait. En effet, un seul morceau de trajectoire était rouge, il n’y avait donc qu’un seul vecteur drift “commun” à toutes les trajectoires.

En bas de la figure H.3, on trouve les résultats de ma méthode pour ces trois mêmes trajectoires. L’algorithme EM converge toujours bien, pour tous les points de départ choisis aléatoirement. Et l’algorithme de Viterbi classe les trajectoires de manière parfaitement conforme à la simulation (figure H.1).

Méthode Monnier :



Ma méthode :



Trajectoires coloriées avec les résultats finaux de l'algorithme EM

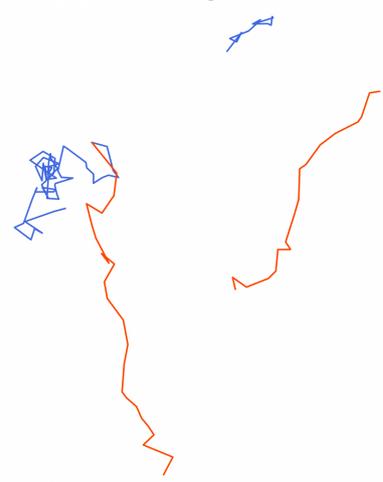


FIGURE H.3 – Résultats de la méthode Monnier (en haut) et de ma méthode (en bas) sur les trajectoires de la Simulation 2.

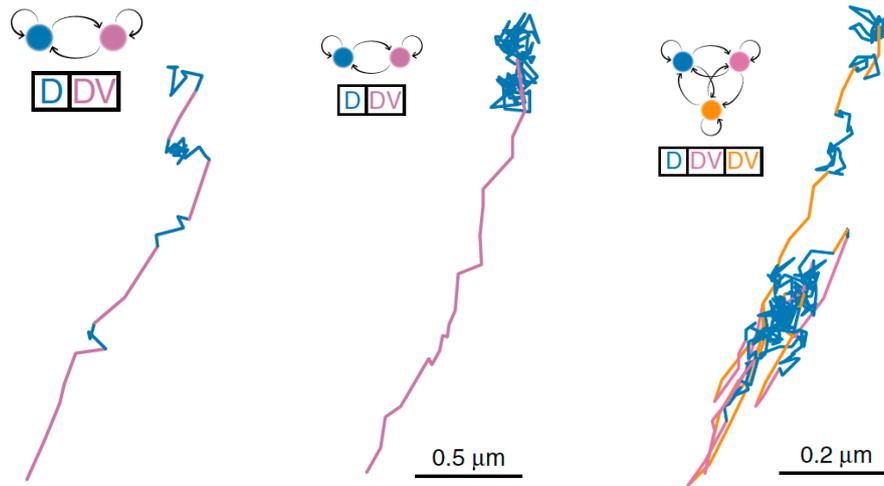


FIGURE H.4 – Trois trajectoires traitées individuellement par la méthode de [MBP⁺15] coloriées avec code couleur : **bleu** pour les morceaux de trajectoires browniens, et les autres couleurs pour les mouvements superdiffusifs.

Malheureusement, pour l'étude d'un grand nombre de trajectoires issues de données réelles, il est difficile de s'imaginer qu'il soit pertinent de supposer que toutes les particules considérées se déplacent avec le même vecteur de drift. Il faudrait donc proposer une adaptation de cet algorithme si l'on souhaite que celui-ci reste efficace pour étudier un pooling de trajectoires.

C'est ce même problème que l'on rencontre si on fait tourner la méthode Monnier sur les deux premières trajectoires présentées en figure H.4 (qui sont les trajectoires présentées dans l'article [MBP⁺15]). Si on étudie les deux premières trajectoires ensemble, on obtient alors les résultats présentés dans la figure H.5 : la combinaison la plus probable est de voir deux mouvements browniens (un représenté en rose et l'autre en bleu).

En revanche, si on fait tourner ma méthode sur ces deux mêmes trajectoires, on obtient les résultats affichés dans la figure H.6.

La figure H.7 affiche les résultats de ma méthode dans le cas où on considère que 3 mouvements sont possibles : un mouvement brownien de coefficient de diffusion σ , un mouvement brownien dirigé de coefficient de diffusion σ et de drift de norme $\|v\|$ et un processus d'Ornstein-Uhlenbeck de coefficient de diffusion σ et de force de rappel λ .

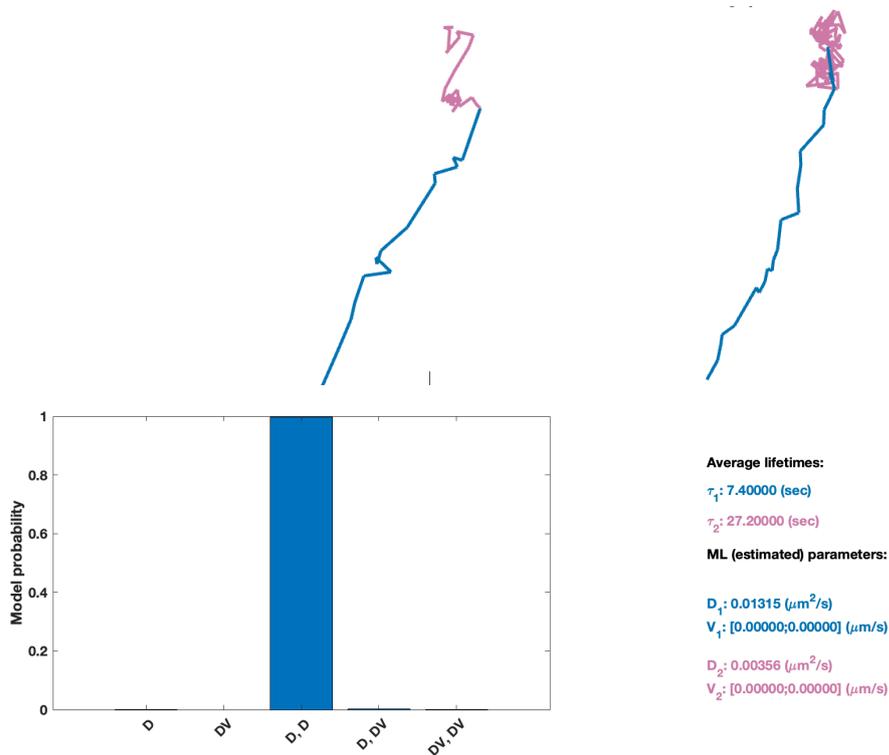


FIGURE H.5 – Deux premières trajectoires de la figure H.4 traitées ensemble par la méthode de [MBP⁺15] coloriées avec code couleur : bleu pour les morceaux qui suivent un premier mouvement brownien, et rose pour les morceaux qui suivent un deuxième mouvement brownien.

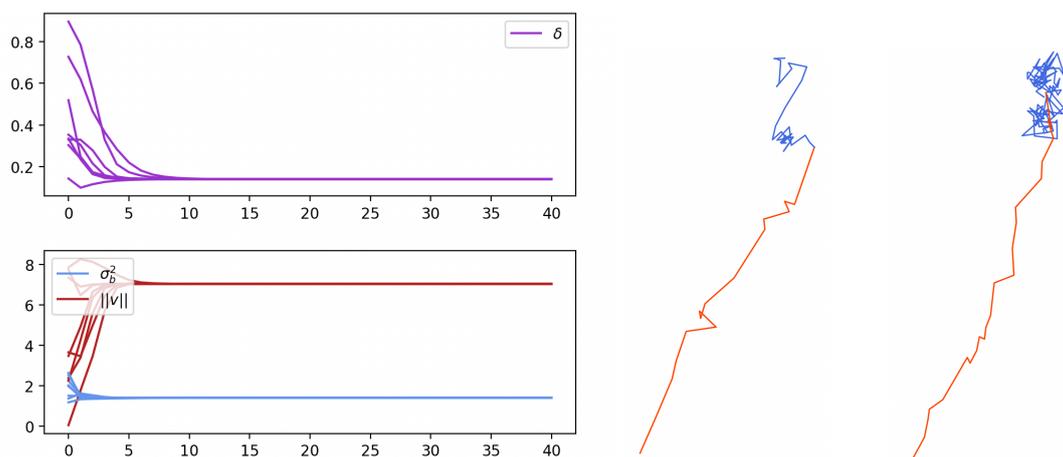


FIGURE H.6 – Deux trajectoires traitées par ma méthode selon le modèle “2 mouvements : un brownien et un brownien dirigé”. À gauche : convergence de l’algorithme EM et à droite : résultat de la classification des morceaux de trajectoires par l’algorithme de Viterbi, coloriées avec le code couleur : **bleu** pour les morceaux de trajectoires browniens, **rouge** pour les mouvements browniens dirigé

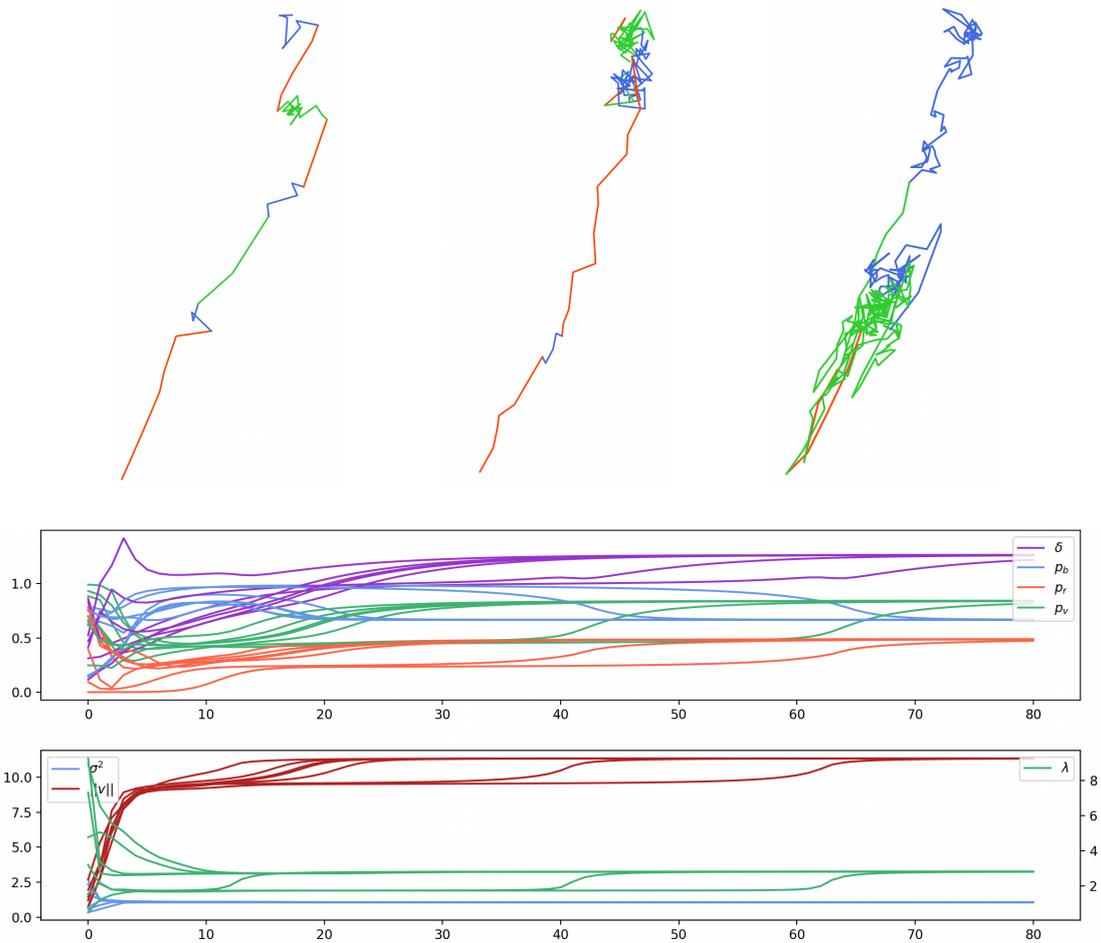


FIGURE H.7 – Trois trajectoires traitées par ma méthode selon le modèle “trois mouvements avec même σ ”. En bas : convergence de l’algorithme EM pour plusieurs points de départ choisis au hasard et en haut : résultat de la classification des morceaux de trajectoires par l’algorithme de Viterbi, coloriées avec le code couleur : **bleu** pour les morceaux de trajectoires browniens, **rouge** pour les mouvements browniens dirigés et **vert** pour ceux qui suivent un processus d’Ornstein-Uhlenbeck.

Conclusion

Dans ce chapitre conclusif, nous reviendrons sur les limites des différentes études que nous avons menées au cours de cette thèse. Cela nous amènera ensuite à exhiber plusieurs perspectives envisageables pour la suite.

1 Limites

1.1 Erreurs de pré-traitement

L'une des principales limitations de notre étude réside dans le fait que nous ne prenons pas en compte les erreurs potentielles pouvant survenir lors du pré-traitement des données réelles. En effet, notre analyse commence à partir de données déjà pré-traitées, que nous considérons comme étant véridiques et exemptes d'erreurs. Cependant, dans un contexte de traitement de données réelles, il est possible que des erreurs se glissent durant cette phase initiale, telles que des détections incorrectes, des absences de détection, ou des erreurs de suivi des particules. Ces erreurs peuvent significativement affecter la qualité des données sur lesquelles repose notre modèle. Par exemple, une fausse détection pourrait entraîner une mauvaise interprétation de la dynamique des particules, tandis qu'une erreur de suivi pourrait brouiller les trajectoires observées.

Dans le cadre du modèle BDMM, de telles erreurs ne sont pas prises en compte, et il pourrait être pertinent de considérer l'introduction d'un bruit artificiel dans le système afin de modéliser et quantifier l'impact de ces imprécisions. Ce bruit pourrait permettre de simuler les erreurs de pré-traitement et ainsi offrir une approche plus robuste, capable de traiter des données légèrement erronées sans altérer la validité des conclusions.

1.2 Limites computationnelles

L'algorithme EM, tel que présenté dans le chapitre 5 et appliqué dans le chapitre 6, constitue un autre point de limitation de notre étude en raison de sa lourdeur computationnelle. Cet algorithme, bien qu'efficace pour estimer les paramètres manquants dans des modèles statistiques, est gourmand en calcul et nécessite un temps conséquent avant d'atteindre la convergence. Pour donner un ordre d'idée, dans le cadre de notre étude sur les données réelles, présentée dans le chapitre 6, le processus de convergence prend plusieurs heures.

Une autre limite concerne la facilité d'utilisation de l'algorithme de simulation du processus BDMM. En effet, une amélioration de l'interface de cet algorithme pourrait le rendre "clé en main" et facilement accessible à tout scientifique souhaitant étudier un système de particules en mouvement, sans avoir besoin de compétences élaborées en langage Python. Cette généralisation pourrait permettre d'élargir la portée du modèle et de l'utiliser comme un outil de modélisation et d'analyse plus universel, capable de s'adapter à différents contextes biologiques.

2 Perspectives

2.1 Étude d'autres problèmes biologiques

Le modèle BDMM que nous avons développé présente un cadre très général et flexible, ce qui le rend adaptable à un large éventail de phénomènes biologiques. En particulier, avec l'apport de nouvelles données issues de divers domaines de la biologie, il serait intéressant de l'appliquer à des problématiques différentes de celles abordées dans cette étude. Par exemple, le modèle BDMM pourrait permettre d'étudier la dynamique d'autres molécules impliquées dans l'exocytose, ou encore de modéliser le trafic intracellulaire d'autres particules telles que les bactéries ou les virus.

Au delà des applications biologiques, il serait aussi intéressant de tirer profit du modèle BDMM dans d'autres domaines, comme en écologie pour l'étude de la dynamique de plantes ou de populations animale.

2.2 Amélioration du suivi de particules

Le modèle BDMM pourrait également être utilisé pour améliorer les outils de suivi des particules, notamment dans le domaine de la microscopie ou des expériences de suivi de trajectoires à l'échelle moléculaire. Ces outils se basent souvent sur des modèles de déplacement a priori pour interpréter les données et fournir des résultats de suivi précis. Le BDMM pourrait jouer un rôle important en tant que modèle a priori, en offrant une représentation plus réaliste et plus flexible des trajectoires suivies par les particules.

En intégrant ce modèle dans les outils de suivi actuels, il serait possible de réduire les erreurs d'estimation des trajectoires, en particulier dans les situations où

les particules changent fréquemment de régime de mouvement ou évoluent dans des environnements complexes. Cela pourrait améliorer la précision des analyses et fournir des informations plus détaillées sur les dynamiques sous-jacentes des systèmes biologiques étudiés.

2.3 Propriétés théoriques du modèle BDMM

Pour que les résultats théoriques obtenus avec le modèle BDMM soient plus en accord avec les observations faites sur des données réelles, certaines propriétés théoriques du modèle (propriété LAN et son corollaire démontrés au chapitre 3) pourraient être adaptées dans le cas d'observations en temps discret plutôt que continu.

Ensuite, en laissant le pas de discrétisation tendre vers zéro, on pourrait théoriquement retrouver les résultats obtenus dans un cadre de temps continu, offrant ainsi une passerelle entre théorie et pratique.

2.4 Analyse des trajectoires

L'analyse des trajectoires, notamment la détection de changements de dynamique au sein d'une même trajectoire de particule, constitue une autre direction de recherche prometteuse. En complément de l'étude menée dans le chapitre 6, il existe plusieurs axes à explorer pour approfondir cette étude.

2.4.1 Plusieurs coefficients de dérive par trajectoire

Une première amélioration serait d'envisager que chaque segment de trajectoire, qui suit un mouvement de type 2 (brownien dirigé), puisse être orienté dans une direction différente. En effet, dans l'étude actuelle, nous faisons l'hypothèse qu'un seul coefficient de dérive s'applique à l'ensemble de la trajectoire, ce qui contraint une particule qui effectue un mouvement dirigé dans la direction η , puis change de régime, à reprendre sa trajectoire dirigée dans la même direction initiale η (si un nouveau changement de régime s'opère). Une approche plus réaliste consisterait à permettre plusieurs directions de dérive pour une même trajectoire, ce qui refléterait mieux la variabilité des trajectoires observées dans les systèmes biologiques.

Une extension de l'étude présentée considérant un nombre N de directions de dérive par trajectoire devrait pouvoir se réaliser en ajoutant de nouvelles inconnues

au système. En revanche, la complexité de l'estimation du nombre de direction de dérive à considérer pour chacune des trajectoires pourrait poser des difficultés et constituer l'objet d'une nouvelle étude approfondie.

2.4.2 Normes de dérive variables

Une autre piste consisterait à adapter le modèle pour prendre en compte des normes de dérive variables selon les trajectoires. En d'autres termes, chaque trajectoire pourrait avoir sa propre norme de dérive, ce qui ajouterait un degré de liberté au modèle. Toutefois, cela soulève des questions d'identifiabilité, car il pourrait devenir difficile de distinguer un mouvement brownien avec un grand coefficient de diffusion d'un mouvement brownien dirigé avec un coefficient de diffusion élevé mais une petite norme de dérive.

2.4.3 Étude d'autres types de mouvement

Une extension naturelle de l'étude actuelle serait d'inclure d'autres types de mouvements que les trois analysés ici. Nous avons déjà exploré cette idée en remplaçant le mouvement de type 3 (mouvement confiné) par un deuxième mouvement brownien. Les résultats obtenus étaient également satisfaisants, bien que nous ayons finalement choisi d'opter pour un processus d'Ornstein-Uhlenbeck, qui semble biologiquement mieux représenter le comportement confiné des particules.

On pourrait aussi étudier des cas où les particules suivent des mouvements non markoviens. Cela ne poserait pas de problème au niveau de la simulation, mais soulèverait des questions théoriques complexes.

2.4.4 Prise en compte de plusieurs sauts dans l'intervalle de discrétisation

Actuellement, toute la Partie III de notre étude repose sur l'hypothèse que les observations sont discrètes et qu'un seul saut peut se produire entre deux observations. Il serait possible de quantifier l'erreur induite par cette hypothèse en s'appuyant sur des majorations connues de l'erreur en fonction du pas de discrétisation. On pourrait imaginer de refaire cette étude dans le cas où plusieurs sauts peuvent avoir lieu entre deux points d'observation, mais cela demanderait sûrement beaucoup d'investissement.

2.4.5 Validation expérimentale

Enfin, une validation expérimentale de notre méthode sur des données simulées pourrait être envisagée. Il serait intéressant de comparer le nombre de points de trajectoire dont le mouvement est correctement estimé par notre méthode par rapport à d'autres méthodes d'analyse de dynamique de trajectoires. Cette comparaison pourrait ensuite être étendue aux données réelles, afin d'évaluer les performances de notre approche dans un cadre expérimental concret.

3 Conclusion

Cette thèse s'inscrit dans un cadre d'interdisciplinarité, visant à apporter une réponse mathématique à une problématique issue de la biologie. Pour atteindre cet objectif, nous avons intégré et développé des outils issus de la statistique, pour s'ajuster au mieux aux besoins de la biologie, en tenant compte des particularités des systèmes observés.

Le modèle BDMM que nous avons introduit a permis de modéliser et analyser la dynamique d'un système complexe de protéines. En particulier, il a permis de décrire en détail les interactions spatio-temporelles entre les protéines. Ce modèle a apporté un cadre théorique permettant d'anticiper des variations dans le comportement des protéines en fonction de différents paramètres, renforçant ainsi notre compréhension des dynamiques sous-jacentes.

Cette collaboration scientifique, mêlant la rigueur des mathématiques et l'expérimentation en biologie, s'est avérée fructueuse. Elle a permis non seulement d'enrichir la recherche biologique par des outils plus précis et prédictifs, mais aussi d'établir un dialogue interdisciplinaire où chaque domaine a contribué à affiner les méthodes de l'autre. Les outils développés au cours de cette thèse ouvrent des perspectives intéressantes pour l'étude de systèmes microbiologiques complexes, offrant ainsi de nouvelles voies d'exploration pour les futures recherches dans ce domaine.

BIBLIOGRAPHIE

- [AABA10] Steven S. Andrews, Nathan J. Addy, Roger Brent, and Adam P. Arkin. Detailed simulations of cell biology with smoldyn 2.1. *PLOS Computational Biology*, 6(3) :1–10, 03 2010.
- [AN12] Krishna B Athreya and Peter E Ney. *Branching processes*, volume 196. Springer Science & Business Media, 2012.
- [ASP⁺19] Marloes Arts, Ihor Smal, Maarten W Paul, Claire Wyman, and Erik Meijering. Particle mobility analysis using deep learning and the moment scaling spectrum. *Scientific reports*, 9(1) :17160, 2019.
- [BAD⁺21] Anaïs Badoual, Misa Arizono, Audrey Denizot, Mathieu Ducros, Hugues Berry, U Valentin Nägerl, and Charles Kervrann. Simulation of astrocytic calcium dynamics in lattice light sheet microscopy images. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 135–139. IEEE, 2021.
- [BGM⁺14] Jérôme Boulanger, Charles Gueudry, Daniel Münch, Bertrand Cinquin, Perrine Paul-Gilloteaux, Sabine Bardin, Christophe Guérin, Fabrice Senger, Laurent Blanchoin, and Jean Salamero. Fast high-resolution 3d total internal reflection fluorescence microscopy by incidence angle scanning and azimuthal averaging. *Proceedings of the National Academy of Sciences*, 111(48) :17164–17169, 2014.
- [BKB09] Jérôme Boulanger, Charles Kervrann, and Patrick Bouthemy. A simulation and estimation framework for intracellular dynamics and trafficking in video-microscopy and fluorescence imagery. *Medical image analysis*, 13(1) :132–142, 2009.
- [BKV18] Vincent Briane, Charles Kervrann, and Myriam Vimond. Statistical analysis of particle trajectories in living cells. *Phys Rev E.*, 97((6-1) :062121), 2018.

- [BL24] Lisa Balsollier and Frédéric Lavancier. Parametric estimation and LAN property of the birth-death-move process with mutations. *arXiv :2404.19367*, 2024.
- [BLSK23] Lisa Balsollier, Frédéric Lavancier, Jean Salamero, and Charles Kervrann. A generative model to simulate spatiotemporal dynamics of biomolecules in cells. *Biological Imaging*, 3 :e22, 2023.
- [BM15] Vincent Bansaye and Sylvie Méléard. *Stochastic models for structured populations*, volume 16. Springer, 2015.
- [BN13] Paul C Bressloff and Jay M Newby. Stochastic models of intracellular transport. *Reviews of Modern Physics*, 85(1) :135, 2013.
- [Bou23] Dominique Bourgeois. Single molecule imaging simulations with advanced fluorophore photophysics. *Communications Biology*, 6(1) :53, 2023.
- [BVK19] Vincent Briane, Myriam Vimond, and Charles Kervrann. An overview of diffusion models for intracellular dynamics analysis. *Briefings in Bioinformatics*, 2019. bbz052.
- [CAV⁺10] Inhee Chung, Robert Akita, Richard Vandlen, Derek Toomre, Joseph Schlessinger, and Ira Mellman. Spatial control of egf receptor activation by reversible dimerization on living cells. *Nature*, 464(7289) :783–787, 2010.
- [CSDC⁺14] Nicolas Chenouard, Ihor Smal, Fabrice De Chaumont, Martin Mavska, Ivo F Sbalzarini, Yuanhao Gong, Janick Cardinale, Craig Carthel, Stefano Coraluppi, Mark Winter, et al. Objective comparison of particle tracking methods. *Nature methods*, 11(3) :281–289, 2014.
- [FG07] Torben Fattler and Martin Grothaus. Strong feller properties for distorted brownian motion with reflecting boundary condition and an application to continuous n-particle systems with singular interactions. *Journal of Functional Analysis*, 246(2) :217–241, 2007.

- [GCK⁺18] Sanjana Gupta, Jacob Czech, Robert Kuczewski, Thomas M Bartol, Terrence J Sejnowski, Robin EC Lee, and James R Faeder. Spatial stochastic modeling with mcell and cellblender. *arXiv preprint arXiv :1810.00499*, 2018.
- [GFSM20] Ángel F. García-Fernández, Lennart Svensson, and Mark R. Morelande. Multiple target tracking based on sets of trajectories. *IEEE Transactions on Aerospace and Electronic Systems*, 56(3) :1685–1707, 2020.
- [GLGW13] Naama Gal, Diana Lechtman-Goldstein, and Daphne Weihs. Particle tracking in living cells : a review of the mean square displacement method and beyond. *Rheologica Acta*, 52 :425–443, 2013.
- [HH17] Nathanaël Hozé and David Holcman. Statistical methods for large ensembles of super-resolution stochastic single particle trajectories in cell biology. *Annual Review of Statistics and Its Application*, 4 :189–223, 2017.
- [HL00] Reinhard Hopfner and Eva Löcherbach. *Limit theorems for null recurrent markov processes*. Univ., Fachbereich Mathematik, 2000.
- [Ibr13] Has’ Minskii Ibragimov. *Statistical estimation : asymptotic theory*, volume 16. Springer Science & Business Media, 2013.
- [JLM⁺08] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature methods*, 5(8) :695, 2008.
- [Ker22] Charles Kervrann. Biomolecule trafficking and network tomography-based simulations. In *Biomedical Image Synthesis and Simulation*, pages 543–569. Elsevier, 2022.
- [Kle12] Fima C Klebaner. *Introduction to stochastic calculus with applications*. World Scientific Publishing Company, 2012.

- [KM57] Samuel Karlin and James McGregor. The classification of birth and death processes. *Transactions of the American Mathematical Society*, 86(2) :366–400, 1957.
- [KSA⁺15] Charles Kervrann, Carlos Óscar Sánchez Sorzano, Scott T Acton, Jean-Christophe Olivo-Marin, and Michael Unser. A guided tour of selected image processing and analysis methods for fluorescence and electron microscopy. *IEEE Journal of Selected Topics in Signal Processing*, 10(1) :6–30, 2015.
- [Kut04] Yury A Kutoyants. *Statistical inference for ergodic diffusion processes*. Springer Science & Business Media, 2004.
- [KWFA18] Nickolay Korabel, Thomas A. Waigh, Sergei Fedotov, and Viki J. Allan. Non-markovian intracellular transport with sub-diffusion and run-length dependent detachment rate. *PLOS ONE*, 13(11) :1–23, 11 2018.
- [LCB⁺20] Matthieu Lagardère, Ingrid Chamma, Emmanuel Bouilhol, Macha Nikolski, and Olivier Thoumine. Fluosim : simulator of single molecule dynamics for fluorescence live-cell and super-resolution imaging of membrane proteins. *Scientific Reports*, 10(1) :1–14, 2020.
- [LDH09] Thibault Lagache, Emmanuel Dauty, and David Holcman. Quantitative analysis of virus and plasmid trafficking in cells. *Physical Review E*, 79(1) :011921, 2009.
- [LLG21] Frédéric Lavancier and Ronan Le Guével. Spatial birth-death-move processes : basic properties and estimation of their intensity functions. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 83(4) :798–825, 2021.
- [LLGM22] Frédéric Lavancier, Ronan Le Guével, and Emilien Manent. Feller and ergodic properties of jump-move processes with applications to interacting particle systems. *to appear in Journal of Applied Probability*, *arXiv :2204.02851*, 2022.

- [Löc02a] Eva Löcherbach. Lan and lamn for systems of interacting diffusions with branching and immigration. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 38(1) :59–90, 2002.
- [Löc02b] Eva Löcherbach. Likelihood ratio processes for markovian particle systems with killing and jumps. *Statistical inference for stochastic processes*, 5(2) :153–177, 2002.
- [LPZK20] Frédéric Lavancier, Thierry Pécot, Liu Zengzhen, and Charles Ker-
vran. Testing independence between two random sets for the ana-
lysis of colocalization in bioimaging. *Biometrics*, 76(1) :36–46, 2020.
- [LS13] Robert S Liptser and Albert N Shiryaev. *Statistics of random pro-
cesses II : Applications*, volume 6. Springer Science & Business Me-
dia, 2013.
- [Lus92] Harald Luschgy. Local asymptotic mixed normality for semimartin-
gale experiments. *Probability theory and related fields*, 92(2) :151–
176, 1992.
- [Man23] Émilien Manent. *Propriétés de Feller, d'ergodicité et inférence
non-paramétrique du processus jump-move*. PhD thesis, Université
Rennes 2; Université de Rennes, 2023.
- [MBP⁺15] Nilah Monnier, Zachary Barry, Hye Yoon Park, Kuan-Chung Su, Za-
chary Katz, Brian P English, Arkajit Dey, Keyao Pan, Iain M Chee-
seman, Robert H Singer, et al. Inferring transient particle transport
dynamics in live cells. *Nature methods*, 12(9) :838–840, 2015.
- [MGVGM⁺21] Gorika Muñoz-Gil, Giovanni Volpe, Miguel Angel Garcia-March,
Erez Aghion, Aykut Argun, Chang Beom Hong, Tom Bland, Stefano
Bo, J Alberto Conejero, Nicolás Firbas, et al. Objective comparison
of methods to decode anomalous diffusion. *Nature communications*,
12(1) :6253, 2021.
- [MH17] Naoki Masuda and Petter Holme. *Temporal network epidemiology*.
Springer, 2017.

- [MS94] Jesper Møller and Michael Sørensen. Statistical analysis of a spatial birth-and-death process model with a view to modelling linear dune fields. *Scandinavian journal of statistics*, 21(1) :1–19, 1994.
- [Øks13] Bernt Øksendal. *Stochastic differential equations : an introduction with applications*. Springer Science & Business Media, 2013.
- [PBB⁺14] Thierry Pécot, Patrick Bouthemy, Jérôme Boulanger, Anatole Chesel, Sabine Bardin, Jean Salamero, and Charles Kervrann. Background fluorescence estimation and vesicle segmentation in live cell imaging with conditional random fields. *IEEE Transactions on Image Processing*, 24(2) :667–680, 2014.
- [PG19] Arne Pommerening and Pavel Grabarnik. *Individual-based methods in forest ecology and management*. Springer, 2019.
- [PKB⁺08] Thierry Pécot, Charles Kervrann, Sabine Bardin, Bruno Goud, and Jean Salamero. Patch-based markov models for event detection in fluorescence bioimaging. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 95–103. Springer, 2008.
- [PLUE13] Fredrik Persson, Martin Lindén, Cecilia Unoson, and Johan Elf. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nature methods*, 10(3) :265–269, 2013.
- [Pre75] Chris Preston. Spatial birth and death processes. *Advances in applied probability*, 7(3) :371–391, 1975.
- [PZB⁺18] Thierry Pécot, Liu Zengzhen, Jérôme Boulanger, Jean Salamero, and Charles Kervrann. A quantitative approach for analyzing the spatio-temporal distribution of 3d intracellular events in fluorescence microscopy. *eLife*, 7 :e32311, 2018.
- [QSE91] Hong Qian, Michael P Sheetz, and Elliot L Elson. Single particle tracking. analysis of diffusion and flow in two-dimensional systems. *Biophysical journal*, 60(4) :910–921, 1991.

- [RDJ⁺17] Philippe Roudot, Liya Ding, Khuloud Jaqaman, Charles Kervrann, and Gaudenz Danuser. Piecewise-stationary motion modeling and iterative smoothing to track heterogeneous particle motions in dense environments. *IEEE Transactions on Image Processing*, 26(11) :5395–5410, 2017.
- [RS01] Eric Renshaw and Aila Särkkä. Gibbs point processes for studying the development of spatial-temporal stochastic processes. *Computational statistics & data analysis*, 36(1) :85–105, 2001.
- [Sad19] Yukio Sadahiro. Analysis of the appearance and disappearance of point objects over time. *International Journal of Geographical Information Science*, 33(2) :215–239, 2019.
- [SFS⁺97] James Schaff, Charles C Fink, Boris Slepchenko, John H Carson, and Leslie M Loew. A general computational framework for modeling cellular structure and function. *Biophysical journal*, 73(3) :1135–1146, 1997.
- [SPB⁺19] Daniel Sage, Thanh-An Pham, Hazen Babcock, Tomas Lukes, Thomas Pengo, Jerry Chao, Ramraj Velmurugan, Alex Herbert, Anurag Agrawal, Silvia Colabrese, et al. Super-resolution fight club : assessment of 2d and 3d single-molecule localization microscopy software. *Nature methods*, 16(5) :387–395, 2019.
- [SX08] Dominic Schuhmacher and Aihua Xia. A new metric between distributions of point processes. *Advances in applied probability*, 40(3) :651–672, 2008.
- [VHHK16] Varun Venkataramani, Frank Herrmannsdörfer, Mike Heilemann, and Thomas Kuner. Suresim : simulating localization microscopy experiments from ground truth models. *Nature methods*, 13(4) :319–321, 2016.
- [VVP14] Ba-Ngu Vo, Ba-Tuong Vo, and Dinh Phung. Labeled random finite sets and the bayes multi-target tracking filter. *IEEE Transactions on Signal Processing*, 62(24) :6554–6567, 2014.

Titre : Simulation, estimation paramétrique et classification des trajectoires d'un processus birth-death-move avec mutations.

Mot clés : Processus birth-death-move, estimation paramétrique, propriété LAN, simulation et synthèse d'images, dynamique intracellulaire et mouvement moléculaire, modèles de Markov cachés, algorithme EM.

Résumé : Cette thèse se concentre sur un processus statistique, dit "birth-death-move" avec mutations (BDMM), qui permet de modéliser la dynamique d'un système de particules qui se déplacent au fil du temps, tandis que de nouvelles particules peuvent apparaître et que certaines particules existantes peuvent disparaître. Ce modèle prend en compte les effets spatiaux et les interactions entre les particules, ce qui le rend adapté pour comprendre la dynamique et la coordination des particules à l'intérieur des cellules. Il offre ainsi un cadre robuste pour étudier une gamme variée de phénomènes biologiques. Après avoir

présenté le processus, des exemples d'applications en biologie et un outil pour le simuler, nous nous intéressons à la propriété de normalité asymptotique locale du modèle paramétrique. Pour finir, nous présenterons une méthode de classification des trajectoires d'un processus BDMM en fonction de leur mouvement, dans le cas particulier où ces dernières sont indépendantes. Tout au long de cette thèse, les différentes études seront illustrées par leur application sur un jeu de données réelles présentant la dynamique des protéines Langerin et Rab-11 au cours du phénomène d'exocytose.

Title: Simulation, parametric estimation and trajectories' classification of a birth-death-move process with mutations.

Keywords: Birth-death-move process, parametric estimation, LAN property, simulation and image synthesis, intracellular dynamics and molecular motion, Hidden Markov Models, EM algorithm.

Abstract: This thesis focuses on a statistical process, known as "birth-death-move" with mutations (BDMM), which models the dynamics of a system of particles that move over time, while new particles may appear and some existing particles may disappear. This model takes into account spatial effects and interactions between particles, making it suitable for understanding the dynamics and coordination of particles within cells. It thus offers a robust framework for studying a wide range of biological phenomena. After presenting the

process, examples of applications in biology and a tool for simulating it, we turn to the property of local asymptotic normality of the parametric model. Finally, we present a method for classifying the trajectories of a BDMM process according to their tracking motion, in the particular case where the trajectories are independent. Throughout this thesis, the various studies will be illustrated by their application to a real data set presenting the dynamics of the Langerin and Rab-11 proteins during exocytosis.